UNIT 1

SPREAD SPECTRUM MODULATION

INTRODUCTION, FREQUENCY HOPPING MULTIPLE ACCESS

Q.1. Write an introductory note on spread spectrum modulation.

Or

What do you mean by spread spectrum technique ? Explain.

(R.G.P.V., June 2016)

Ans. Spread spectrum (SS) techniques were originated to meet the requires of military and commercial communications. This technique is a wideband modulation technique. In spread spectrum, the bandwidth expansion factor is more. In reality, this bandwidth expansion factor does not combat additive white Gaussian noise (AWGN), as does wideband frequency modulation (FM). Since the spread spectrum technique is not useful in combating white noise.

One way of classifying the spread spectrum is defined as follows –

(i) In averaging system, the reduction of interference occurs due to the interference can be averaged over a large time interval.

(ii) In avoidance system, the interference reduction takes place due to the signal is made to eliminate the interference over a large time fraction.

As far as classification by modulation technique is concerned, two main techniques employed in spread spectrum systems are frequency hopping (FH) and direct sequence (DS).

Q.2. Explain the basic concept of spread spectrum and discuss the basic techniques of spread spectrum. ***(R.G.P.V., Dec. 2016)***

Ans. Fig. 1.1 shows the basic principle of a spread spectrum system. Input is applied into a channel encoder that produces an analog signal with a relatively narrow bandwidth around some center frequency. This signal is further modulated using a sequence of digits known as a spreading code or spreading sequence. Typically, but not always, the spreading code is produced by a pseudonoise or pseudorandom number generator. The effect of this modulation is to increase significantly the bandwidth (spread the spectrum) of the signal to be transmitted. On the receiving end, the same digit sequence is used to demodulate the spread spectrum signal. Finally, the signal is fed into a channel decoder to recover the data.

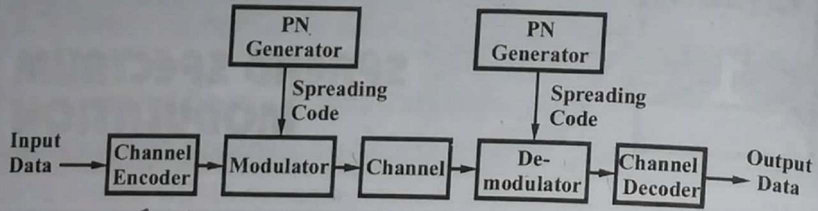


Fig. 1.1 General Model of Spread Spectrum System

Also refer to the ans. of Q.1.

Q.3. *Write down the goals of spread spectrum systems.*

Or

Give some potential advantages of spread spectrum system. ***(R.G.P.V., June 2016)***

Ans. Spread spectrum systems also meet the following goals as –

(i) Antijam capability

(ii) Secure communications

(iii) Operation with a low-energy spectral density

(iv) Multipath protection

(v) Ranging

(vi) Multiple-access capability without external control

Q.4. *Write down the advantages of spread spectrum transmission.*

Or

List the advantages of spread spectrum. ***(R.G.P.V., Dec. 2016)***

Ans. The advantages of spread spectrum transmission are as follows –

(i) Spread spectrum signals can be overlaid onto bands where other systems are already operating, with minimal performance impact to both systems.

(ii) Spread spectrum is a wideband signal that has a superior performance over traditional radios on frequency selective fading multipath channel. Spread spectrum provides a robust and reliable transmission in urban and indoor environments where wireless transmission suffers from heavy multipath conditions.

(iii) The anti-interference characteristics of spread spectrum are important in some applications, such as networks operating on manufacturing floors, where some signal interference environment can be harsh.

(iv) Cellular systems designed with CDMA spread spectrum technology provide greater operational flexibility and overall system capacity than systems built on FDMA or TDMA access methods.

(v) The convenience of unlicensed spread spectrum operation in ISM bands in the United States is attractive to manufacturers and users a like.

Q.5. *Give the disadvantages of spread spectrum transmission.*

Ans. The disadvantages of spread spectrum transmission are as follows –

(i) One disadvantage is the increased complexity of receivers that have to despread a signal. Today despreading can be performed upto high data rates thanks to digital signal processing.

(ii) Another problem is the large frequency band that is needed due to the spreading of the signal. Although spread signals appear more like noise, they still raise the background noise level and may interfere with other transmissions if no special precautions are taken.

Q.6. *What is jamming margin ?* ***(R.G.P.V., Dec. 2016)***

Ans. The capability of a system to perform in interfering (hostile) environment, is expressed by ***jamming margin.*** Jamming margin takes into account the requirement for a useful system output signal to noise ratio (minimum $(S/N)_{out}$ of the system) and permits for internal losses. Hence, the jamming margin is calculated by the formula –

$$M_j = \text{Jamming margin} = G_p - [L_{sys} + (S/N)_{min\ out}]$$

where, L_{sys} = System implementation losses.

As an example, a system with 30 dB process gain, minimum $(S/N)_{out}$ of 10 dB and L_{sys} of 2 dB would have an 18 dB jamming margin (M_j). Thus, it could not be expected to operate with interference more than 18 dB above the desired signal. Remember that jamming is not always the result of an intentional act. Sometimes, the jamming signal is caused by natural phenomena and sometimes it is the result of self-interference caused by the multipath, in which delayed versions of the signal, arriving via alternative paths, interfere with the direct path transmission. However, it should also be noted that, though the spread spectrum technique provides good process gain or unwanted signal rejection, it does not provide highest process gain for every situation. Thus, to obtain optimum results, the spread spectrum techniques should be combined with other signal to noise ratio improving techniques.

Q.7. *Discuss the working principle of frequency hopping in brief.*

Ans. The frequency hopping (FH) technique is used to change the frequency of carrier signal of a narrowband transmission system therefore the transmission is completed in one band of frequency only for a short while.

The spreading factor is the ratio between the narrowband transmission bandwidth and the bandwidth over which the frequency of carrier signal is hopped.

The signal is only for a short while in the vulnerable band by changing the frequency of carrier signal frequently. The frequency hopping pattern has to be predictable to the desired receiver, but unpredictable for the enemy, making them unable to follow the frequency hopping.

Normally, frequency hopping causes to a whitening of characteristics of received signal. There is implicit averaging over attenuation of channel. In addition, a different interferer is active at each frequency of carrier signal, such that frequency hopping also causes to an averaging over all interferers. Since it decreases probability of disastrous scenarios, this is advantageous for several types of receivers and hence reduces the needed link margins.

Q.8. Discuss the various types of frequency hopping.

Ans. There are two types of frequency hopping and there is a considerable difference in performance for both the system as follows –

(i) ***Slow Frequency Hopping*** – This frequency hopping sends one or more symbols on each frequency. Frequently, slow frequency hopping can be used in conjunction with time division multiple access i.e., each timeslot is sent on a given frequency of carrier signal and the next time slot then changes to a different frequency of carrier signal. In this case, the additional effort for synchronization is small, since the receiver has to synchronize for the next time slot anyway. For effective frequency hopping, interleaving and coding has to distribute information assigning to one bit of source over many timeslots.

(ii) ***Fast Frequency Hopping*** – This frequency hopping changes the frequency of carrier signal many times during transmission of one symbol. Alternatively, the transmission of each separate symbol can be spread over a large bandwidth. Therefore, the interference or fading effects are combated for each symbol separately. It follows from elementary Fourier considerations that transmission of each part of a symbol needs large bandwidth than that of a narrowband system. In addition, different contributions combining belonging to one symbol has to use processing that operates faster than at the rate of symbol.

Q.9. Discuss the principle working of frequency hopping multiple access.
(R.G.P.V., June 2017)

Or

What is the concept of frequency hopping multiple access technique in spread spectrum modulation ?
(R.G.P.V., June 2015)

Ans. Frequency hopping is used as multiple access technique which is as spectrally efficient as frequency division multiple access (FDMA) and time division multiple access (TDMA). In these consideration, we can distinguish between synchronized and unsynchronized systems as follows –

(i) ***Synchronized Case*** – In this case, the base station (BS) radiates the signal to all mobile stations (MSs) at the same time. An example with three available bands (carrier frequencies) is shown in fig. 1.2. Obviously, the base station can send to three mobile stations simultaneously during one time interval. For two mobile stations (A and B), the base station sends signal to mobile station A in band 2 during the first interval of time. At the same time interval, band 3 is free, so the base station can send to mobile station B in that band. During the second interval of time, the base station now sends to mobile station A in band 1, and mobile station B in band 2. During the third interval of time, mobile station A is serviced in band 3, and mobile station B in band 1. Then the whole sequence repeats. The signals for all the mobile stations use the same sequence of hopping and there is never a collision between them. Hence, obviously, we have the same capacity as FDMA, with the added frequency diversity benefit.

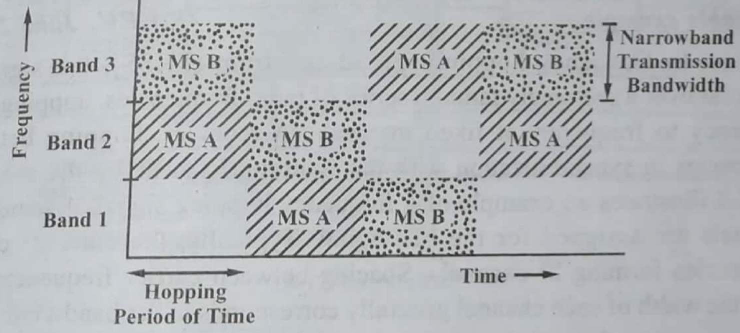


Fig. 1.2 Frequency Hopping Multiple Access with Synchronized Mobile Stations

(ii) ***Unsynchronized Case*** – This case can either take place in simple networks where timing advance is unpredictable, for consideration of ad-hoc networks, or in intercell interference. It should be noted that, any delay between the signals of different mobile stations is possible because of the lack of synchronization, including a zero-delay. When all mobile stations use the same sequence of hopping, then such a zero-delay causes to catastrophic collisions, in which different mobile stations interfere with each other all the time. In order to remove this problem, different hopping sequences are used for each mobile stations as shown in fig. 1.3. These hopping sequences are designed in such a manner that during each hopping cycle, the time duration of exactly

one timeslot is disturbed, whereas the remainder time is guaranteed to be collision free. Clearly, the performance of unsynchronized system is worse than that of a synchronized system.

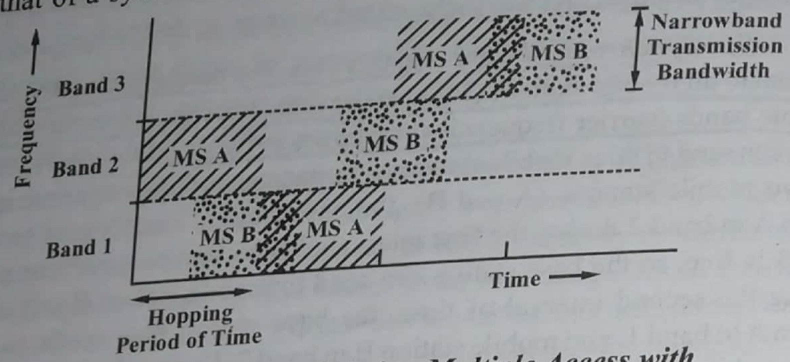


Fig. 1.3 Frequency Hopping Multiple Access with Unsynchronized Mobile Stations

Q.10. With the help of block diagram explain the working of frequenc hopped spread spectrum technique. **(R.G.P.V., June 2014)**

Or

Explain the frequency hopped spread spectrum technique with the hel of suitable example. **(R.V.P.V., June 2016)**

Ans. In frequency hopping spread spectrum (FHSS), the signal i broadcast over a seemingly random series of radio frequencies, hopping from frequency to frequency at fixed intervals. A receiver, hopping between frequencies in synchronization with the transmitter, picks up the message. Fig. 1.4 illustrates an example of a frequency hopping signal. A number of channels are assigned for the FH signal. Typically, there are 2^k carrie frequencies forming 2^k channels. Spacing between carrier frequencies and thus, the width of each channel generally corresponds to the bandwidth of the

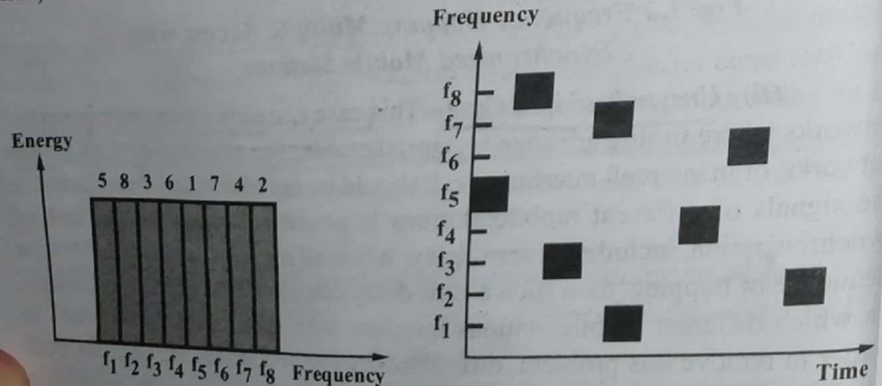
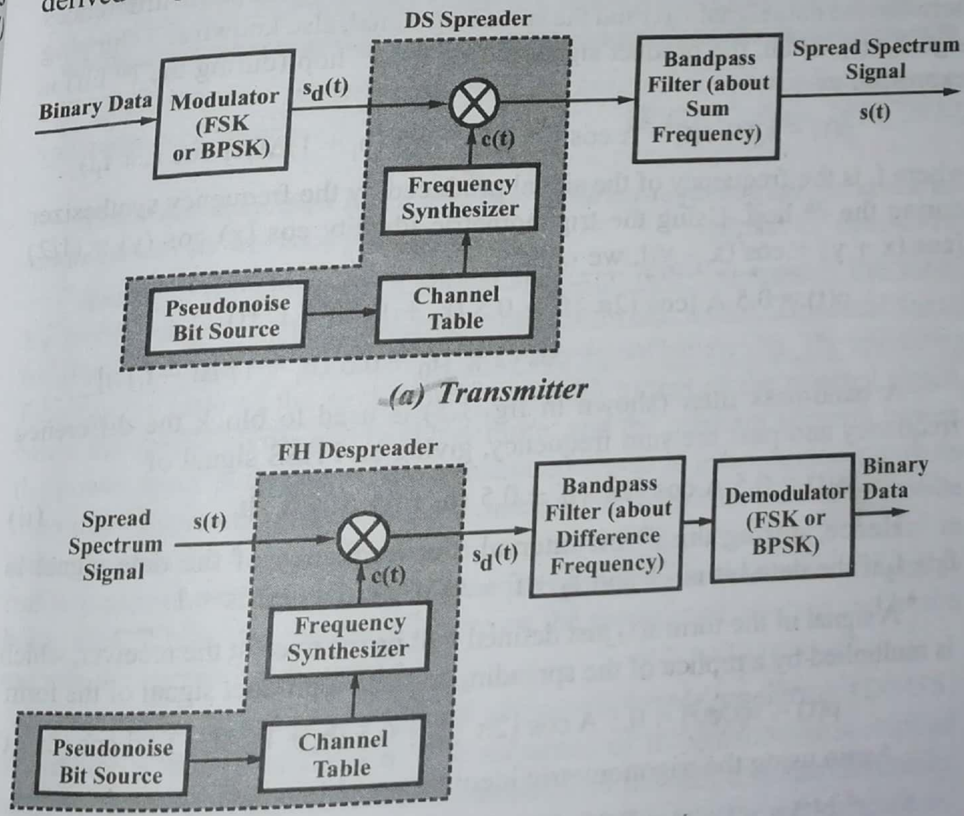


Fig. 1.4 Example of Frequency Hopping

input signal. The transmitter operates in one channel at a time for a fixed interval. As an example, the IEEE 802.11 standard uses a 300 m-sec interval. In this interval, some number of bits are sent using some encoding scheme. A spreading code dictates the sequence of channels. Both transmitter and receiver use the same code to tune into a sequence of channels in synchronization.

Fig. 1.5 shows a typical block diagram for a frequency hopping system. For transmission, binary data are given into a modulator using some digital to analog encoding scheme, like frequency shift keying (FSK) or binary phase shift keying (BPSK). Resulting signal is centered around some base frequency. A pseudonoise (PN) or pseudorandom number, source serves as an index into a table of frequencies; this is the spreading code. Each k bits of the PN source specifies one of the 2^k carrier frequencies. At each successive interval (each k PN bits), a new carrier frequency is chosen. Then, this frequency is modulated by the signal generated from the initial modulator to generate a new signal with the same shape but now centered on the chosen carrier frequency. On reception, the spread spectrum signal is demodulated using the same sequence of PN-derived frequencies and then demodulated to generate the output data.



(a) Transmitter

(b) Receiver

Fig. 1.5 Block Diagram of Frequency Hopping Spread Spectrum System

Fig. 1.5 shows that the two signals are multiplied. The working of thin approach, using BFSK as the data modulation scheme, is given as follows. The FSK input to the FHSS system is defined as –

$$s_d(t) = A \cos[2\pi (f_0 + 0.5(b_i + 1) \Delta f)t], \text{ for } iT < t < (i + 1) T \quad ...(i)$$

where, A = Amplitude of the signal

f_0 = Base frequency

b_i = Value of the i^{th} bit of data (+1 for binary 1 and –1 for binary 0)

Δf = Frequency separation

T = Bit duration; data rate = $1/T$.

Hence, during the i^{th} bit interval, the frequency of the data signal is f_0 if the data bit is –1 and $f_0 + \Delta f$ is the data bit is +1.

Frequency synthesizer produces a constant frequency tone whose frequency hops among a set of 2^k frequencies, with the hopping pattern found by k bits from the PN sequence. For convenience, suppose that the duration of one hop is the same as the duration of one bit and we ignore phase differences between the data signal $s_d(t)$ and the spreading signal, also known as a chipping signal, c(t). Then, the product signal during the i^{th} hop (during the i^{th} bit) is expressed as –

$$p(t) = s_d(t)\, c(t) = A \cos [2\pi \{f_0 + 0.5 (b_i + 1)\Delta f\}t] \cos (2\pi f_i t)$$

where f_i is the frequency of the signal produced by the frequency synthesizer during the i^{th} hop. Using the trigonometric identity $\cos (x) \cos (y) = (1/2) [\cos (x + y) + \cos (x - y)]$, we obtain –

$$p(t) = 0.5 A [\cos (2\pi \{f_0 + 0.5 (b_i + 1) \Delta f + f_i \}t)$$
$$+ \cos (2 \pi \{f_0 + 0.5 (b_i + 1) \Delta f - f_i\}t)]$$

A band-pass filter (shown in fig. 1.5) is used to block the difference frequency and pass the sum frequency, giving an FHSS signal of

$$s(t) = 0.5 A \cos [2\pi \{f_0 + 0.5 (b_i + 1) \Delta f + f_i\}]t \quad ...(ii)$$

Hence, during the i^{th} bit interval, the frequency of the data signal is $f_0 + f_i$ if the data bit is –1 and $f_0 + f_i + \Delta f$ if the data bit is + 1.

A signal of the form s(t) just defined will be received at the receiver, which is multiplied by a replica of the spreading to yield a product signal of the form

$$p(t) = s(t)c(t) = 0.5 A \cos [2\pi \{f_0 + 0.5 (b_i + 1) \Delta f + f_i\}t] \cos (2\pi f_i t)$$

Again using the trigonometric identity, we obtain –

$$p(t) = s(t)c(t) = 0.25 A [\cos \{2\pi (f_0 + 0.5 (b_i + 1) \Delta f + f_i + f_i)\}t$$
$$+ \cos \{2\pi (f_0 + 0.5 (b_i + 1) \Delta f)\}t]$$

A band-pass filter, shown in fig. 1.5, is used to block the sum frequency and pass the difference frequency, giving an signal of the form of $s_d(t)$, defined by the equation (i) –

$$0.25 A \cos [2\pi \{f_0 + 0.5 (b_i + 1) \Delta f\}t]$$

Another common modulation technique used in conjunction with FHSS is multiple FSK (MFSK). To encode the digital input L bits at a time, MFSK uses $M = 2^L$ different frequencies. Transmitted signal is of the form –

$$s_i(t) = A \cos 2 \pi f_i t, \qquad 1 \le i \le M$$

where, M = Number of different signals = 2^L

L = Number of bits per signal element

$f_i = f_c + (2i - 1 - M)f_d$

f_c = Carrier frequency

f_d = Difference frequency.

CDMA, CELLULAR CDMA SYSTEMS

Q.11. Discuss the principle and working of direct sequence spread spectrum technique. ***(R.G.P.V., June 2016, 2017)***

Or

Draw the simplified block diagram of direct sequence spread spectrum and describe its working principle. ***(R.G.P.V., Dec. 2016)***

Ans. The direct-sequence spread spectrum (DS-SS) spreads the signal by multiplying the send signal with large bandwidth signal. This total signal bandwidth is equal to the wideband spreading signal bandwidth. The spreading factor is the ratio of the new signal bandwidth to that of the original signal. Since the spread signal bandwidth is more, and the constant transmit power, the power spectral density of the transmitted-signal is less i.e., depending on the value of spreading factor and the distance between base station or mobile station, it can lie below the noise power spectral density. This is useful in military communications, because enemies cannot achieve whether a signal is being sent. While, desired users can invert the spreading operation and hence recover the narrowband signal.

The block diagram of a direct sequence spread spectrum (DS-SS) transmitter is shown in fig. 1.6. The sequence of information is multiplied with a broadband signal which was obtained by modulating a sinusoidal carrier signal with a spreading sequence. On the other side, this can be obtained as multiplying each information symbol of duration T_S with a spreading sequence p(t) before the process of modulation. We consider that the spreading sequence is M_C chips long, in which each chip has the duration $T_C = T_S/M_C$. Since the bandwidth is the inverse of the chip duration, the total signal bandwidth is –

$$W = \frac{1}{T_C} = \frac{M_C}{T_S}$$

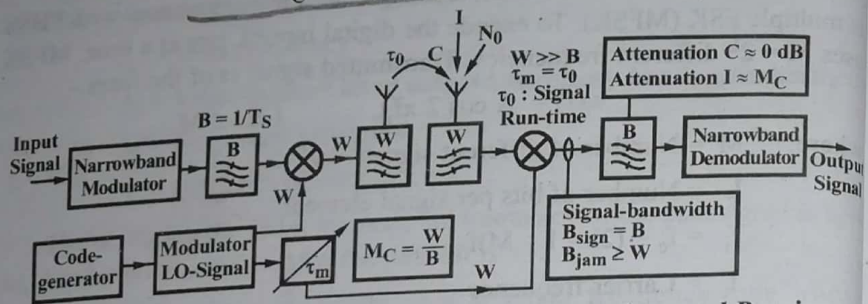


Fig. 1.6 Direct-sequence Spread Spectrum Transmitter and Receiver

In the receiver, the spreading operation is inverted by correlating the received signal with the sequence of spreading i.e., multiply by the time-inverted complex-conjugate sequence of spreading, and integrate over appropriate interval. Therefore after correlation, this technique reverses bandwidth spreading, the desired signal again has a bandwidth of $1/T_S$. In this desired signal, the received signal also comprises noise, other interferers of wideband signal, and possibly interferers of narrowband signal. The despreading operation does not affects the effective noise bandwidth and wideband interferers in a significant manner. On the other hand, narrowband interferers are actually spread over a bandwidth W. The despreading of signal passes by a LPF (lowpass filter) of bandwidth $B = 1/T_S$. This leaves the desired signal essentially unchanged, but the noise power, wideband interferers, and narrowband interferers are decreased by a factor $1/M_C$. Hence, the direct-sequence spread spectrum (DS-SS) has the same SNR (signal-to-noise ratio) as a narrowband system at the symbol demodulator. For a narrowband system, the power of noise is N_0/T_S at the demodulator. For a direct-sequence spread spectrum (DS-SS) system, the power of noise is $N_0/T_C = N_0M_C/T_S$, at the input of receiver, which is decreased by narrowband filtering. Hence, the noise power is N_0/T_S at the input of detector.

Q.12. *Explain process gain in spread spectrum system.*

Or

Why is there a process gain in a spread spectrum system ?

(R.G.P.V., Dec. 2016)

Ans. Process gain is defined as the difference between the output and input signal to noise ratios. This is a main parameter, which characterizes the properties of spread spectrum communication systems. As an example, in the frequency modulation case, it is referred to as FM improvement and is equal to 10 log $(3/2)m^2$, where m is the FM modulation index given by $m = \Delta\omega/\omega_m$, $\Delta\omega$ being the frequency deviation. Hence, for a given processor with an input S/N ratio of 10 dB and output S/N ratio of 16 dB would have a process gain of 6 dB. Practically, the process gain is given by the following formula as a rule of thumb, namely –

$$\text{Process gain } (G_p) = \frac{BW_{rf}}{R_{inf}} \qquad ...(i)$$

where, BW_{rf} = RF bandwidth of the transmitted spectrum signal

R_{inf} = Information rate or data rate in the information base band channel.

Importance of process gain can be further understood in the following way. It is well known that a digital communication system operates with a given required probability of error when the ratio of the energy per bit to the spectral density of the noise (or interference) exceeds some threshold value, typically around 10 dB. The received signal power is given by the product of the energy per bit and the bit rate. Hence,

$$P_r = E_b(\text{J/bit})\ R(\text{bit/sec}) \qquad ...(ii)$$

Interference power from the other spread spectrum users is the product of their combined interference spectral density N_0 and the receiver input bandwidth B or W_{ss}.

$$P_i\,(\text{watts}) = N_0\ (\text{watts/Hz}).\ W_{ss}(\text{Hz}) \qquad ...(iii)$$

Thus, the signal to interference power ratio at the input of the receiver is given as –

$$SIR = \frac{P_r}{P_i} = \frac{E_b}{N_0} \cdot \frac{R}{W_{ss}} \qquad ...(iv)$$

However, in a spread spectrum signal, W_{ss} is much greater as compared to R and therefore, E_b/N_0 could be in the acceptable range (around 10) even when SIR is very small.

As an example, if $E_b/N_0 = 10$ and $W_{ss}/R = 1000$, SIR = 0.01 or –20 dB. Due to this reason, W_{ss}/R is known as processing gain of the system. Hence, it is possible for a spread spectrum receiver to operate even when its input signal is literally 'buried' in noise or interference.

Q.13. *With the help of block diagram explain the working of direct sequence spread spectrum technique. Also explain processing gain.*

(R.G.P.V., June 2014)

Ans. Refer to the ans. of Q.11 and Q.12.

Q.14. Give the principle and mathematical representation of direct sequence spread spectrum.

Ans. For military communications, the direct sequence (DS) spreading principle itself (i.e., multiplication through the wideband signal) can be seen as a modulation technique. The code division multiple access is used to obtain multiaccess capability. Each user is allocated a distinct spreading code, which calculates the wideband signal i.e., multiplication with the information symbols. At the receiver, the signals from other users, which were modulated through a distinct spreading sequence, appear as additional noise. Hence, several users can send in a wideband simultaneously.

The desired signal can be achieved by correlating the received signal with the spreading signal of the desired user at the receiver. Hence, the other users become wideband interferers. After passing by the despreader, the interference power observe through the detector is the cross correlation function (CCF) between the spreading sequence of the desired user and the spreading sequence of the interfering user. Hence, the CCF in ideal case is,

$$CCF_{j,k}(t) = 0 \text{ for } j \neq k \quad \text{...(i)}$$

For all user j and k. Alternatively, the orthogonal code sequences are needed. Better orthogonality is obtained for at most M_C spreading sequences. This is immediately observed through the fact that M_C orthogonal sequences span an M_C-dimensional space, and any other sequence of that duration is shown as a linear combination.

If the sequences of spreading are not orthogonal, then the receiver gets finite interference suppression by a factor ACF/CCF. This suppression factor is M_C if the many spreading sequences are shifted PN-sequences.

Mathematical Representation – In this representation, the binary phase shift keying (BPSK) modulation is used and that better synchronization is available between sender and receiver. Then, we define four signal components at the receiver as follows –

(i) User Signal – We consider $c_{i,k}$ is the i-th information symbol of the k-th user, and $r_{i,k}$ the corresponding receive signal. Let the interchip interference, intersymbol interference and noise are zero-mean processes, then the desired received signal is equal to the transmit symbol as follows –

$$E\{r_{i,k}|c_{i,k}\} = \sqrt{(E_C)_k}\, c_{i,k} \int_{-\infty}^{\infty} |H_R(f)|^2\, df \quad \text{...(ii)}$$

where $(E_C)_k$ denotes the chip energy of the k-th user, and $H_R(f)$ the transfer function of the receive filter normalized to $\int_{-\infty}^{\infty} |H_R(f)|^2\, df = 1$.

(ii) Interchip Interference – If the filter of receiver has a finite-duration impulse response, then the chip convolution with this impulse response lasts longer than the chip itself. Hence, the signal provides interchip interference after the receive filter. If the spreading sequences are zero-mean, then the interchip interference raises the received signal variance as follows –

$$(E_C)_k \sum_{i \neq 0} \left[\int_{-\infty}^{\infty} \cos(2\pi \text{ if } T_C) |H_R(f)|^2\, df \right]^2 \quad \text{...(iii)}$$

(iii) Noise raises the variance by $N_0/2$.

(iv) Co-channel Interference – Let the mean of the received signal is unchanged through the co-channel interference, as the transmitted chips of interfering users are not dependent of the data symbols and desired users chips. The co-channel interference raises the variance by –

$$\sum_{j \neq k} \frac{(E_C)_j}{2T_C} \int_{-\infty}^{\infty} |H_R(f)|^4\, df \quad \text{...(iv)}$$

Q.15. Describe code division multiple access (CDMA).

(R.G.P.V., May 2018)

Ans. CDMA is a multiplexing scheme used with spread spectrum. The technique works in the following way. We begin with a data signal with rate D, which is called the bit data rate. Each bit is broken into k chips according to a fixed pattern which is specific to each user, referred to as the user's code.

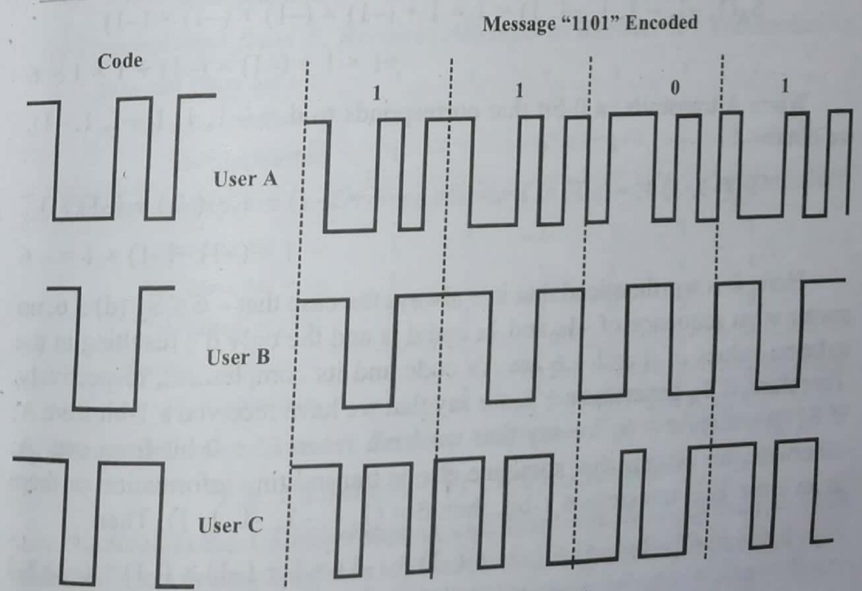


Fig. 1.7 Example of CDMA

New channel has a chip data rate of kD chips per second. To understand, we consider a simple example with k = 6. It is simplest to characterize a code as a sequence of 1s and –1s. Fig. 1.7 depicts the codes for three users A, B and C, each of which is communicating with the same base station receiver R. Hence, the code for user A is

$C_A = \langle 1, -1, -1, 1, -1, 1 \rangle$. Likewise, user B has code

$C_B = \langle 1, 1, -1, -1, 1, 1 \rangle$ and user C has $C_C = \langle 1, 1, -1, 1, 1, -1 \rangle$.

Now, we take the case of user A communicating with the base station. It is assumed that the base station known A's code. For convenience, we assume that communication is already synchronized so that the base station known when to look for codes. If A wants to transmit a 1-bit, A sends its code as a chip pattern $\langle 1, -1, -1, 1, -1, 1 \rangle$. If a 0-bit is to be transmitted, A sends the complement (1s and –1s reversed) of its code, $\langle -1, 1, 1, -1, 1, -1 \rangle$. At the base station, receiver decodes the chip patterns. In case the receiver R receives a chip pattern $d = \langle d1, d2, d3, d4, d5, d6 \rangle$ and the receiver is seeking to communicate with a user u so that it has at hand u's code, $\langle c1, c2, c3, c4, c5, c6 \rangle$, the receiver performs electronically the following decoding function –

$$S_u(d) = d1 \times c1 + d2 \times c2 + d3 \times c3 + d4 \times c4 + d5 \times c5 + d6 \times c6$$

Subscript u on S simply denotes that u is the user. Let us consider that the user u is in fact A. When A transmits a 1-bit, then d is $\langle 1, -1, -1, 1, -1, 1 \rangle$ and the preceding computation using S_A becomes –

$$S_A(1, -1, -1, 1, -1, 1) = 1 \times 1 + (-1) \times (-1) + (-1) \times (-1) + 1 \times 1 + (-1) \times (-1) + 1 \times 1 = 6$$

When A transmits a 0-bit that corresponds to $d = \langle -1, 1, 1, -1, 1, -1 \rangle$, we obtain –

$$S_A(-1, 1, 1, -1, 1, -1) = -1 \times 1 + 1 \times (-1) + 1 \times (-1) + (-1) \times 1 + 1 \times (-1) + (-1) \times 1 = -6$$

Here, it is worthnoticed that it is always the case that $-6 \le S_A(d) \le 6$, no matter what sequence of –1s and 1s and d is and the only d's resulting in the extreme values of 6 and – 6 are A's code and its complement, respectively. Therefore, if S_A generates a + 6, we say that we have received a 1-bit from A; if S_A generates a – 6, we say that we have received a 0-bit from user A; otherwise, we assume that someone else is transmitting information or there is an error. If B transmits a 1-bit, then $d = \langle 1, 1, -1, -1, 1, 1 \rangle$. Then,

$$S_A(1, 1, -1, -1, 1, 1) = (-1) \times (-1) + (-1) \times 1 + (-1) \times (-1) + (-1) \times 1 + (-1) \times (-1) + (-1) \times 1 = 0$$

Hence, the undesired signal (from B) does not show up at all. It can be easily verified that if B had transmitted a 0-bit, the decoder would generate a value of 0 for S_A again. That is, if the decoder is linear and if A and B send signals s_A and s_B, respectively, at the same time, then $S_A(s_A + s_B) = S_A(s_A) + S_A(s_B) = S_A(s_A)$ because the decoder ignores B when it is using A's code. Codes of A and B that have the property that $S_A(C_B) = S_B(C_A) = 0$ are known as orthogonal. These codes are very nice to have but there are not all that many of them. Fig. 1.8 shows the summary of the example from the preceding discussion.

User A	1	–1	–1	1	–1	1
User B	1	1	–1	–1	1	1
User C	1	1	–1	1	1	–1

(a) User's Codes

Transmit (data bit = 1)	1	–1	–1	1	–1	1	
Receiver codeword	1	–1	–1	1	–1	1	
Multiplication	1	1	1	1	1	1	= 6
Transmit (data bit = 0)	–1	1	1	–1	1	–1	
Receiver codeword	1	–1	–1	1	–1	1	
Multiplication	–1	–1	–1	–1	–1	–1	= –6

(b) Transmission from A

Transmit (data bit = 1)	1	1	–1	–1	1	1	
Receiver codeword	1	–1	–1	1	–1	1	
Multiplication	1	–1	1	–1	–1	1	= 0

(c) Transmission from B, Receiver Attempts to Recover A's Transmission

Transmit (data bit = 1)	1	1	–1	1	1	–1	
Receiver codeword	1	1	–1	–1	1	1	
Multiplication	1	1	1	–1	1	–1	= 2

(d) Transmission from C, Receiver Attempts to Recover B's Transmission

B (data bit = 1)	1	1	–1	–1	1	1	
C (data bit = 1)	1	1	–1	1	1	–1	
Combined signal	2	2	–2	0	2	0	
Receiver codeword	1	1	–1	1	1	1	
Multiplication	2	2	2	0	2	0	= 8

(e) Transmission from B and C, Receiver Attempts to Recover B's Transmission

Fig. 1.8 Summary of CDMA Example

Practically, the CDMA receiver can filter out the contribution from undesired users or they appear as low-level noise. However, if there are several users competing for the channel with the user the receiver is trying to listen to, or if the signal power of one or more competing signals is very high perhaps since it is very near the receiver, the system breaks down.

Q.16. *Write various properties of maximal-length sequences.*
(R.G.P.V., May 2018)

Ans. The various properties of maximal-length sequences are as follows.

(i) Run Length – Among the runs of ones and zeros in every period it is desirable that about one half of the runs of every type are of length 1, one fourth are of length 2, one eighth are of length 3 and so on. Assume a PN code is 000100110101111.

Number of runs = 8

$$\underbrace{000}_{3}\ \underbrace{1}_{1}\ \underbrace{00}_{2}\ \underbrace{11}_{2}\ \underbrace{0}_{1}\ \underbrace{1}_{1}\ \underbrace{0}_{1}\ \underbrace{1111}_{4}$$

(ii) Autocorrelation – For a maximal length sequence, autocorrelation function is periodic and binary valued. Autocorrelation function states as –

$R_a(\tau) = \dfrac{1}{N_c}$ [Number of agreements – Number of disagreements in comparison of one full period]

```
0 0 0 1 0 0 1 1 0 1 0 1 1 1 1
1 0 0 0 1 0 0 1 1 0 1 0 1 1 1
d a a d d a d a d d d d a a a
```

$$R_a(\tau) = -\frac{1}{15}$$

Q.17. *Discuss the effects of multipath propagation on CDMA in brief.*

Ans. The frequency selectivity effect on a CDMA system is known by seeing at the impulse response of the concatenation spreader-channel-despreader. If the channel is slowly time-variant, then the effective impulse response is defined as follows –

$$h_{eff}(t_i, \tau) = \tilde{p}(\tau) * h(t_i, \tau) \qquad \text{...(i)}$$

where the effective system impulse response $\tilde{p}(\tau)$ denotes the convolution of the transmit and receive spreading sequence as follows –

$$\tilde{p}(\tau) = p_{TX}(\tau) * p_{RX}(\tau)$$
$$= ACF(\tau) \qquad \text{...(ii)}$$

In an ideal spreading sequence, the despreader output provides multiple peaks. Rather, one for each MPC (multi-path component) which is resolved through the receiver i.e., spaced at least T_C apart. Each of the peaks comprises the transmit signal information. Hence, all the peaks must be used in the process of detection. Just using the largest correlation peak would mean that we discard a lot of the arriving signal. A receiver that can use multiple correlation peaks is the Rake receiver, which collects the energy from different multi-path components. A Rake receiver comprises of a bank of correlators. Each correlator is sampled at a distinct time, and hence collects energy from the multi-path component with delay τ. The values of sample from the correlators are then combined and weighted.

On the other hand, the Rake receiver is used as a tapped line, in which outputs are combined and weighted. The tap weights and tap delays both are adjustable and matched to the channel. Normally, the taps can be spaced at least one chip duration apart, but there is no needed for the taps to be spaced at regular intervals. The combination of the Rake receiver and receiver filter establishes a filter which is matched to the receive signal. The Rake receiver is matched to the channel, whereas the receive filter is matched to the transmit signal.

Independent of this explanation, the receiver combines the signal from the several Rake fingers in a coherent manner. Since these signals correspond to several multi-path components, their fading is statistically not dependent. Alternatively, they give delay diversity. Hence Rake receiver is a diversity receiver, and all mathematical techniques remain valid for the diversity treatment.

Q.18. *Discuss the synchronization problem in CDMA system in brief.*

Ans. Synchronization is an estimation problem, where we obtain the optimum sampling time out of an infinitely large ensemble of possible values i.e., the continuous time. Implementation is facilitated by dividing the problem into two partial problems as follows –

(i) Acquisition Problem – A first step achieves, where interval of time (i.e., T_C and $T_C/2$) the optimum sampling time lies. This problem is known as hypothesis testing problem. We can test a finite number of hypothesis, each of which supposes that the sampling time is in a certain interval. The hypothesis are tested serial or parallel.

(ii) Tracking Problem – A control loop is used to fine-tune the sampling time to its accurate value as soon as this interval has been obtained.

A special synchronization sequence is considered which is smaller than the spreading sequence used during data transmission for the acquisition phase. This reduces the hypothesis number which have to be tested, and hence reduces the time which has to be spent on synchronization. In addition, the synchronization sequence is used to have especially better autocorrelation properties. The normal spreading used for data communication is employed for the tracking.

The following synchronizations are important for many system design aspects as –

(i) Synchronization between Base Stations – Global positioning system (GPS) synchronizes the base stations with respect to each other. Therefore, each base station needs a GPS receiver and free line of sight (LOS) to many GPS satellites.

(ii) Synchronization within a Cell – The signals sent through base station are always synchronous, as the base station has control over when to send them. In the uplink case, synchronous arrival of the signals at the base station would need that all the mobile stations arrange their timing advance in such a manner that all signals arrive simultaneously. The timing advance would have to be exact within duration of one chip.

Q.19. What do you mean by code families ?

Ans. In the uplink, the spreading sequences most frequently used are PN-sequences, Gold sequences and Kasami sequences. The autocorrelation properties of PN-sequences are good and PN-sequence shifted versions are again valid codewords with very nearly ideal cross-correlation properties. The suitable combination of PN-sequences generates Gold sequences. The auto-correlation functions of Gold sequences can take on three possible values. Both the offpeak values of the auto-correlation function and cross-correlation function can be upper-bounded.

Now, we can distinguish between small (S), large (L), and very large (VL) Kasami sequences. VL-Kasami sequences have the worst cross-correlation function, but there is a very nearly unlimited number of such sequences. S-Kasami sequences have the best cross-correlation functions, but only a rather small number of such sequences exists. The properties of the different code families is shown in table 1.1.

Table 1.1

Sequence	***Maximum CCF per dB***	***Number of Codes***	***Comment***
VL-Kasami	$\approx -3N_{reg}/2+6$	$2^{N_{reg}/2}(2^{N_{reg}}+1)^2$	Very nearly unlimited number
L-Kasami	$\approx -3N_{reg}/2+3$	$2^{N_{reg}/2}(2^{N_{reg}}+1)$	
S-Kasami	$\approx -3N_{reg}/2$	$2^{N_{reg}/2}$	Good cross correlation function of all Kasami sequences
Gold	$\approx -3N_{reg}/2+1.5$	$2^{N_{reg}}+1$	
PN-sequence		$2^{N_{reg}}-1$	Better Auto Correlation Function

Q.20. Give the selection criteria of in term of code families.

Ans. The quality of spreading codes is generally obtained through the following properties –

(i) Autocorrelation Properties – In an ideal case, the auto correlation function ACF(0) must be equal to the number of chips per symbol M_C, and zero at all other instances. For Pseudonoise (PN) sequences ACF(0) is equal to $M_C - 1$ or nearly equal to M_C, and ACF(n) is equal to –1 for n is not equal to zero. For synchronization, the better properties of the auto correlation function are also used. In addition, ACF properties affect interchip interference in a Rake receiver. The correlator output is equal to the sum of the delayed echoes ACFs.

(ii) Cross-correlation Properties – In an ideal case, all codes must be orthogonal to each other, therefore the interference from other users can be suppressed totally. Orthogonality should be satisfied for arbitrary delays between the several users for unsynchronized systems. Bandwidth spreading and users division are performed by many codes. By using this case, we can distinguish between scrambling codes and spreading codes.

(iii) Number of Codes – A code division multiple access system must permit simultaneous communications of as several user as possible. Therefore, a large number of codes has to be available. The number of chips per symbol M_C limits the number of orthogonal codes. Bad properties of cross-correlation have to be achieved, if more codes are needed. In addition, the condition is complicated by the fact that codes in adjacent cells have to be distinct. After all, it is the only codes that distinguish several users. Now, the codes in cell B cannot be orthogonal to the codes in cell, if cell A has M_C users with all orthogonal codes. As a result, intercell interference cannot be suppressed totally. Although, the codes in cell B are selected in such a manner that the interference to users in the original cell becomes noise like. Then, this technique needs code planning rather frequency planning.

Q.21. What do you mean by Walsh-Hadamard codes ?

Ans. Signals belonging to many users are formed totally orthogonal for the downlink, since they all are radiated by the same base station. A Walsh-Hadamard matrices provides a family of codes that all satisfy these requirements. The n + 1-order Hadamard matrix $H^{(n+1)}_{Had}$ in terms of the n^{th} order matrix can be defined as follows –

$$\mathbf{H}^{(n+1)}_{Had} = \begin{pmatrix} \mathbf{H}^{(n)}_{Had} & \mathbf{H}^{(n)}_{Had} \\ \mathbf{H}^{(n)}_{Had} & \overline{\mathbf{H}}^{(n)}_{Had} \end{pmatrix} \qquad \text{...(i)}$$

where $\overline{\mathbf{H}}$ denotes the modulo-2 complement of **H**. The equation of recursiv can be originated as follows –

$$\mathbf{H}_{\text{Had}}^{(1)} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \quad ...(i)$$

The columns of this matrix are all possible Walsh-Hadamard codeword of length 2. We can obtain $\mathbf{H}_{\text{Had}}^{(2)}$ from the recursion equation as follows –

$$\mathbf{H}_{\text{Had}}^{(2)} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \quad ...(ii)$$

The columns of this matrix represent all possible codewords of duration four.

Q.22. Discuss the principle working of cellular CDMA systems. ***(R.G.P.V., June 2017)***

Or

Write a detail note on cellular CDMA systems. ***(R.G.P.V., June 2015)***

Or

Give an introductory note on CDMA. ***(R.G.P.V., June 2016)***

Ans. We define the situation in which spreading codes are not perfect but there is a large number of them. Distinct users can be distinguished by distinct spreading codes. Therefore, each user in the cell provides interference to all other users because user division is imperfect. Although, the interference for each user raises as well as the number of users raises. As a result, until users find the quality too bad to place calls, quality of transmission reduces gradually. Although, the code division multiple access substitutes, a soft limit on the number of users. Consequently, the number of users depends critically on the SINR (signal-to-interference-and-noise-ratio) needed through the receiver.

Large interference restricts from within the same cell as the desired user, and hence is called intracell interference. Total intracell interference is equal to the sum of several independent contributions, and hence acts approximately such as Gaussian noise. Although, it leads effects that are identical to thermal noise. Frequently, it can be defined by noise rise i.e., the raise in effective noise power than the noise alone $(N_0 + I_0)/N_0$.

A main property of a code division multiple access system is that it uses universal frequency reuse. Alternatively, the same band of frequency is used in all cells and users in distinct cells are distinguished by distinct codes only. Generally, the interference can be defined through the codes, which are used in distinct cells.

Especially in the uplink, the merits of CDMA are related to the fact that interference acts very nearly like noise because of various reasons as follows –

(i) Power control is used for all signals reaching at the base station have nearly the same power.

(ii) The number of users is more in each cell.

(iii) Interference signal from neighbouring cells also provide, from a large number of users. Spreading codes are made in such a manner that all users from a neighbouring cell provide nearly the same percentage to the total intercell interference.

Interference power represents almost no fluctuations because of this process of averaging. While, the power control is used for the signal strength from the desired user is always constant. Hence, the signal-to-interference-and-noise-ratio is constant, and no fading margin has to be used in the link budget.

Q.23. Discuss the power control for the uplink case.

Ans. This can be performed by a closed control loop i.e., the mobile station first transmits with a certain power, then the base station tells the mobile station whether the power is high or low, and the mobile station adjusts its power accordingly. The control loop bandwidth has to be selected so that it can compensate for small-scale fading i.e., has to be on the Doppler frequency order. Because of the time variance and noise in the channel estimate, there is a remaining variance in the powers reaching at the base station.

It should be noted that an open control loop cannot be used to compensate for small-scale fading in a frequency domain duplexing system. Although, an open loop is used in conjunction with a closed loop. The open loop compensates for large-scale variations in the channel, that are nearly the same at uplink and downlink frequencies. Then, the closed loop is compensated for small-scale variations.

Q.24. Discuss the power control for the downlink case.

Ans. This is not necessary for code division multiple access (CDMA) to function in the downlink case i.e., all signals from the base station reach at one mobile station with the same power. Although, it is advantageous to still use power control for the total transmit power low. The send power is reduced for all users within a cell through the same value leaves unchanged the ratio of desired signal power to intracell interference i.e., interference from signals intended for other users in the cell. Therefore, the total interference power is reduced to other cells. While, as the SNR should not fall below a threshold level, we cannot reduce signal power arbitrarily. Hence, the downlink power control is used to reduce the total transmit power whereas keeping the BER or SINR level above a threshold level.

It should be noted that the power control of users in adjacent cells does not provide constant power of the intercell interference. In an adjacent cell, a user power is controlled by its own base stations. Therefore, it "observes" a totally distinct channel to the undesired base station, with temporal variation that the desired base station neither understands nor cares about. As a result, intercell interference is variant temporally.

If all users employ the same data rate, then the interference power from all users is similar. User with higher data rates provide large interference power.

Q.25. *Give the various methods for capacity increases in CDMA system.*

Ans. (i) The Quiet Periods During Speech Transmission – During the quiet periods, a signal with a very low data rate or no signal, has to be sent. In code division multiple access system, not sending information causes to a reduce in total sent power, and hence interference in the system. If reducing the interference power, then additional users permits to place calls. Practically, there is a worst case scenario in which all users in a cell are talking simultaneously, but statistically speaking, this is largely not probable, especially when the number of users is more. Hence, pauses in the conversation are employed by CDMA for capacity improvement.

(ii) Flexible Data Rate – In CDMA system, arbitrary data rates are sent by a suitable choice of spreading sequences. Therefore, the flexible data rate permits for good exploitation of the available spectrum for data transmission.

(iii) Soft Capacity – In CDMA system, the capacity can vary from cell to cell. If cell adds large users, it raises interference to other cells. Hence, it can be possible to have some cells with low capacity, and some with higher. In addition, this can change dynamically, because the traffic changes. This concept is called breathing cells.

(iv) Error Correction Coding – This coding has drawback, that the data rate which is to be sent is raised, that reduces spectral efficiency. While, CDMA consciously raises the amount of data to be sent. Hence, it can be possible to comprise error correction coding without reducing spectral efficiency.

Q.26. *With the help of block diagram explain reverse CDMA channel.*
(R.G.P.V., June 2015)

Ans. Reverse traffic channels and access channels comprises the reverse CDMA channel. The reverse channel signal received at the base station cannot use coherant detection because the MS does not install a system time as at the base station. Hence, the modulation characteristics are different for forward and reverse channels respectively. Fig. 1.9. shows the modulation of the reverse channel at point A. Which is 64 ary orthogonal at a data rate of 9600, 4800, 2400, 1200 bps.

At 28800 code symbols/sec., the real burst transmission rate is fixed. This results in a fixed Walsh chip rate of 307.2 thousand chips per second. Four PN chips spreads each Walsh chip. Spreading PN sequence rate is fixed at 1.2288 Mcps.

Convolutional Encoding – In fig. 1.9, at point B with K = 9 (9 register) and rate 1/3 convolutional encoder –

(i) Every code symbol consists of a fixed data rate of 4800 bps on the access channel, and every symbol repeats one time consecutively.

(ii) The full data rate on the reverse traffic channel is 9600 kbps. When data rate is 4800 kbps, then each symbol repeats one time consecutively, and when data rate is 2400 kbps, each symbol repeats three times consecutively. For the data rate of 1200 kbps, each symbol repeats seven times consecutively.

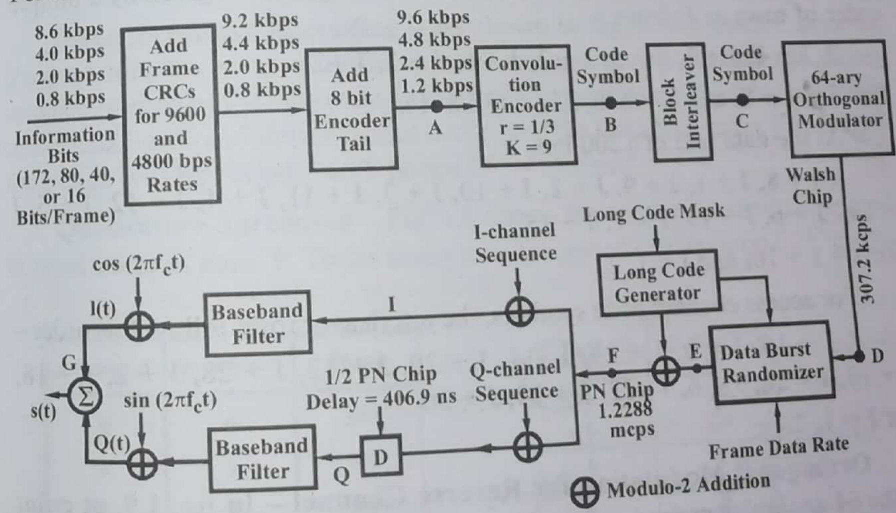


Fig. 1.9 Reverse CDMA Channel Modulation Process

Interleaving – Referring to fig. 1.9 , at point C, the interleaving algorithm will make an array with 32 rows and 18 columns. At 9600 kbps, the interleaver makes a 32 × 18 matrix which is shown in table 1.2 .

Table 1.2 Interleaving Algorithm

Row \ Column	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	33	65	97	129	161	193	225	257	289	321	353	385	417	449	481	513	54
2	2	34	66	99	130	162	194	226	258	290	322	354	386	418	450	482	514	54
3	3	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
4	4	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
5	5	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
6	6	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
7	7	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
32	32	64	96	128	160	192	224	256	288	320	352	384	416	448	480	512	544	576

The transmission sequence is to transmit row by row in a sequence order upto row 32 at 9600 bps. At the data rate of 4800 bps, the transmission sequence is to transmit by the unique order of rows as follows –

Row Number –

1 3 2 4 5 7 6 8 9 11 10 12 13 15 14 16 17 19 18 20 21 23 22 24 25 27 26 28 29 31 30 32

In case of 2400 bps data rate, the transmission sequence is given by a unique order of rows as follows –

J, J + 4, J + 1, J + 5, J + 2, J + 6, J + 3, J + 7

for J = 1 + 8i and i = 0, 1, 2, ..., (32/8–1).

At the data rate of 1200 bps,

J, J + 8, J + 1, J + 9, J + 2, J + 10, J + 3, J + 11, J + 4, J + 12, J + 5, J + 13, J + 6, J + 14, J + 7, J + 15

for J = 1 + 16i and i = 1, 2.

For access channel code symbols, the interleaver rows follow this order–

J, J + 16, J + 8, J + 24, J + 4, J + 20, J + 12, J + 28, J + 2, J + 18, J + 10, J + 26, J + 6, J +22, J + 14, J + 30

for J = 1, 2.

Orthogonal Modulation for Reverse Channel – In fig. 1.9, at point D the 64-ary Walsh codes has 64 codes each 64 bits long. They are orthogonal to each other. Each sixth symbol interpreting each Walsh code of 64 chips is transmitted out. Each 20 ms reverse traffic channel frame will be divided into 16 equal length (that is, 1.25 ms) power control groups numbered from 0 to 15.

20 ms

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15

1.25 ms [12-bits = 36 (12 × 3) Code Symbols
= 36/6 Modulation Symbols
= 1 Power Control Group

Reverse traffic channel and access channel will be direct sequence spread by the long code prior to transmission. Long code will be periodic with period $2^{42} - 1$ chips. This will satisfy the linear recursion specified by the polynomial –

$$p(x) = x^{42} + x^{35} + x^{33} + x^{31} + x^{27} + x^{26} + x^{25} + x^{22} + x^{21} + x^{19} + x^{18} + x^{17} + x^{16} + x^{10} + x^{7} + x^{6} + x^{5} + x^{3} + x^{2} + x^{1} + 1$$

Every PN chip of the long code will be produced by a 42-shift register long-code generator.

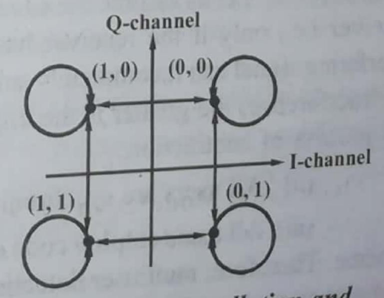
Data Burst Randomizing – As shown in fig. 1.9, at point E the data burst randomizer has produced a masking pattern of 0s and 1s that randomly masks out the redundant data produced by the code repetition. The mask pattern is determined by the data rate of the frame and by a block of 14 bits taken from the long code. These 14 bits will be last 14 bits of the long code used for spreading.

Direct Sequence Spreading – As shown in fig. 1.9, at point F prior to transmission, the reverse traffic channel and the access channel are direct sequence spread by the long code. In this spreading operation, modulo-2 addition of the data burst randomizer output stream and the long code takes place. This long code will be periodic with period $2^{42} - 1$ chips.

Quadrature Spreading – Fig. 1.9 shows the sequences used for spread in quadrature at point F. These sequences are periodic with period 2^{15} chips.

I-channel	Q-channel	Phase
0	0	$\pi/4$
0	1	$-\pi/4$
1	0	$3\pi/4$
1	1	$-3\pi/4$

(a) Reverse CDMA Channel I and Q Mapping



(b) Offset QPSK Constellation and Phase Transition

Fig. 1.10 Reverse CDMA Channel Quadrature Spreading

The spread polynomials of channels I and Q pilot PN sequences are given as

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

$$P_Q(x) = x^{15} + x^{12} + x^{11} + x^{10} + x^6 + x^5 + x^4 + x^3 + \ldots$$

which are of period 2^{15}–1. Fig. 1.10 shows the reverse CDMA channel I and Q mapping for an offset QPSK modulation.

MULTI USER DETECTION, TIME HOPPING IMPULSE RADIO

Q.27. Discuss about the multiuser detection technique.

(R.G.P.V., June 2014)

Or

***Explain in brief multiuser detection technique.* (R.G.P.V., June 2015)**

Ans. Multiuser detection is based on the concept of detecting interference and exploiting the resulting information to serious its effect on the desired signal. The serial interference cancellation is the conceptually simplest version of multiuser detection. We assume a system in which a interfering signal is much higher as compared to the desired signal. This highest signal is first detected and demodulated by the receiver. This signal has a better signal-to-interference-and-noise-ratio (SINR), and hence can be detected without errors. Then its effects are cancelled from the total received signal. Then, the desired signal is obtained at the receiver within the cleaned-up signal. Since this cleaned-up signal comprises only the desired signal and the noise, the SINR is better, and detection process can be performed efficiently.

Q.28. Write down the assumptions of multiuser detection.

Ans. The following assumptions for multiuser detection are given below –

(i) The receiver has complete channel information from the interferer to the receiver. Let us assume the situation in a serial interference cancellation receiver i.e., only if the receiver has complete channel information of the interfering signal can it completely subtract it from the total signal. The higher the interference, the greater is the impact of any channel estimation error on the process of subtraction.

(ii) All users are synchronized.

(iii) All users employ code division multiple access as a multiaccess scheme. Therefore, multiuser detection is also possible for other multiaccess techniques such as TDMA, and can be used, for example is employed to reduce the reuse distance in a TDMA network.

(iv) Only, we use multiuser detection at the receiver.

Q.29. Discuss the linear multiuser detector with the help of block diagram.

Ans. Fig. 1.11 shows the block diagram of a linear multiuser detector system. This system first estimates the signals from several users by despreading using the spreading sequences of the several users. It should be noted that this needs a number of parallel despreaders, each working with a distinct spreading sequence. Then, the outputs from these despreaders are combined linearly. This step of combination is used as filtering using a matrix filter, and used for interference minimization.

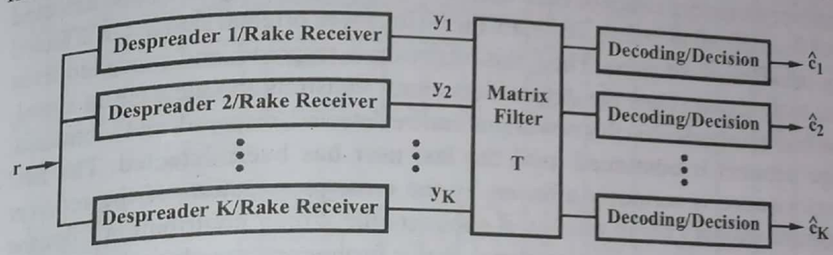


Fig. 1.11 Linear Multiuser Detector System

Decorrelation Receiver – This receiver is the simplest means of multiuser detection. The decorrelation receiver is the equivalent of zero-forcing equalizer. After despreading, the receiver filter output can be defined as follows –

$$\mathbf{y} = \mathbf{n} + \mathbf{cR} \quad \text{...(i)}$$

where **n** denotes the noise vector and the correlation matrix **R** can comprise possible antenna and/or delay diversity. Now, the symbols estimation can be determined simply through filtering with a matrix filter $\mathbf{T} = \mathbf{R}^{-1}$ as follows –

$$\hat{\mathbf{c}} = \mathbf{yR}^{-1} = \mathbf{nR}^{-1} + \mathbf{c} \quad \text{...(ii)}$$

Its simplicity is the merit of this technique and the fact that it is unnecessary to understand the received amplitudes. Only the correlation matrix **R** requires to be obtained. The demerit lies in noise enhancement. The worse the conditioning of the correlation matrix, the large the noise is raised.

Minimum Mean Square Error (MMSE) Receiver – A MMSE multiuser detector provides a balance between noise enhancement and interference suppression. The mean quadratic error $E\{|c-\hat{c}|^2\}$ is a measure for total disturbance. Hence, the matrix filter can be defined as –

$$\mathbf{T} = [\mathbf{I}\sigma_n^2 + \mathbf{R}^{-1}]^{-1} \quad \text{...(iii)}$$

Because of its smaller noise enhancement the distortion of signals can be smaller as compared to the decorrelation receiver, the MMSE does not cause to total interference suppression.

Q.30. Discuss the non-linear multiuser detectors with the help of block diagram.

Ans. The following non-linear multiuser detectors are as follows –

(i) Successive Interference Cancellation (SIC) – This detects user in the sequence of their signal power. Fig. 1.12 shows that each user signal is subtracted from the total signal before the next user is detected. Hence, the successive interference cancellation is a special case of a decision feedback receiver. At receiver, the total signals is received, and it is despread using the distinct spreading codes of each user. Then, the highest signal can be detected and decided on, therefore we obtain the bitstream of original signal, not affected by interference or noise. Then, this bitstream is respread, and cancelled from the total signal. Then the despreaders again transmits the cleaned-up signal, the highest user within this new signal can be detected, respread, and subtracted. The process is continued until the last user has been detected. The SIC performance is seriously affected by the error propagation. If the receiver decides wrongly about one bit, it subtracts the wrong contribution from the total signal, and the remaining signal, that is further processed, not less, suffers from more interference.

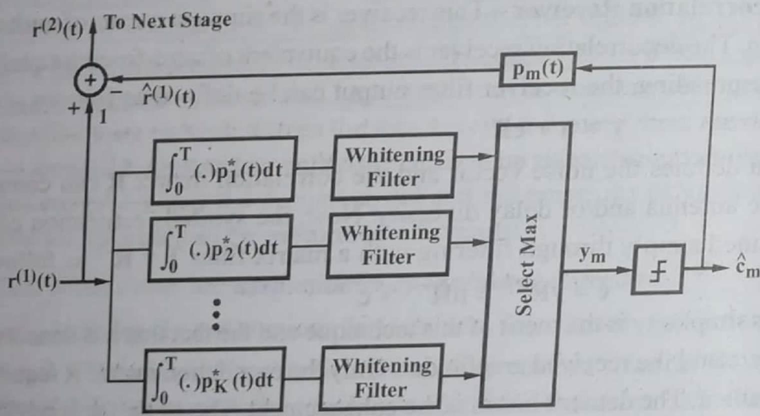


Fig. 1.12 Successive Interference Cancellation

Hard and soft are the two possibilities for interference subtracting. Only a scaled-down version of signal is subtracted for soft subtraction. The interference is totally subtracted for hard subtraction. This does not cause complete signal cancellation, but makes error propagation small of an issue.

(ii) Parallel Interference Cancellation (PIC) – We can also subtract all users simultaneously in place of subtracting interference in a serial fashion. To do this, the first step makes a decision for all users based on, the total received signal. Then, the signals are respread, and contributions from all interferers to the total signal are subtracted. It should be noted that a distinct group of interferers is active for each users. For user 1, user 2.....K represent as interferers and for user 2, user 1 and 3.....K represent as interferers, and so on. Then the canceller next stage uses the "cleaned-up" signals as a basis for a decision, and performs again remodulation and subtraction. This process is continued until a certain number of iterations is reached or until decision no longer change from iteration to iteration as shown in fig. 1.13.

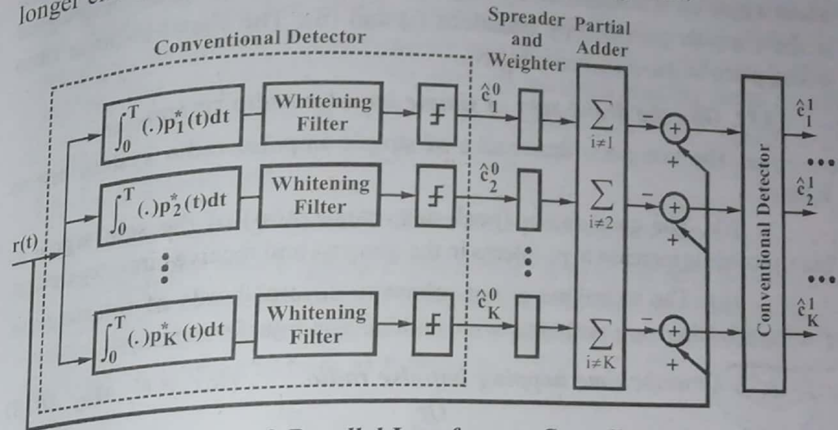


Fig. 1.13 Parallel Interference Cancellation

For parallel interference subtraction, the error propagation may be a adequate problem as well. It should be noted that only the first stage would be required if all decisions were perfect. Although, each signal in the first stage has a worse signal-to-interference ratio, so that there is a high probability of only the highest signal being selected perfectly. Besides, the signal is distorted always by wrong decisions, so that later stages perform even worse than the first.

Q.31. Discuss the simple impulse radio system.

Ans. When the desired spreading bandwidth W is larger, it becomes to spread the spectrum by transmitting short pulses whose amplitude or position comprises the desired information, Frequently, this technique permits the use of simple transmitters and receivers, known as impulse radio.

In simple impulse radio system, a transmitter sending out a single pulse to show one symbol. We consider that the modulation method is orthogonal pulse position modulation for the moment. A pulse can be either transmitted at time $t + T_d$ or at time t, where the duration of a pulse T_d is larger as compared to the T_C. Then, the process of detection in an additive white Gaussian noise channel is exceedingly simple i.e., just we require an energy detector that computes during which of the two possible intervals of time as follows –

$$[t + \tau_{run}, t + \tau_{run} + T_C] \qquad ...(i)$$

or

$$[t + \tau_{run} + T_d, t + \tau_{run} + T_C + T_d] \qquad ...(ii)$$

We obtain more energy. Where, τ_{run} denotes the runtime between transmitter and receiver, which is obtained by the process of synchronization.

Clearly, due to the transmit signal bandwidth is provided by the inverse pulse duration $1/T_C$, a pulsed transmission achieves spreading. The despreading can be performed in a very simple manner i.e., we can make our decision about which bit is transmitted by only recording and using the arriving signal in the intervals provided by equations (i) and (ii). The signal-to-noise ratio taking place at the receiver is E_B/N_0.

Q.32. Give the drawbacks of simple impulse radio system.

Ans. The two main drawbacks of simple impulse radio system are as follows –

(i) The crest factor (peak-to-average ratio) of the sent signal is high, providing increase to problems in the transmit and receive circuitry design.

(ii) The technique is not robust to several kinds of interference. Particularly, there are problems with multiaccess interference.

Q.33. Describe time hopping impulse radio. ***(R.G.P.V., May 2018)***

Or

Discuss the principle of time hopping impulse radio.

(R.G.P.V., June 2017)

Ans. In time-hopping impulse radio (TH-IR), we divide the available symbol time into a number of frames of duration T_f, and send one pulse within each frame is shown in fig. 1.14.

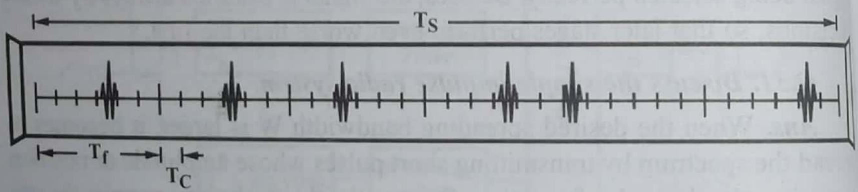


Fig. 1.14 Transmit Waveform of TH-IR for one Symbol, p(t), Showing Chip, Frame and Symbol Duration

Now, the main aim is to vary the pulses position within the frames. Therefore, we transmit the pulse in the third chip of the frame for the first frame, we transmit the pulse in the eighth chip of the frame for the second frame and so on. A Pseudorandom (PN) sequence determines the pulses positions within a frame known as a time-hopping sequence. Mathematically, the transmit signal can be defined as –

$$s(t) = \frac{\sqrt{E_S}}{\sqrt{N_f}} \sum_{i} \sum_{j=1}^{N_f} g(t - jT_f - c_j T_C - b_i T_d - iT_S)$$

$$= \frac{\sqrt{E_S}}{\sqrt{N_f}} \sum_{i} p(t - b_i T_d - iT_S) \qquad \text{...(i)}$$

where N_f denotes the number of frames indicating one information symbol of length T_S, g(t) denotes the sent unit energy pulse of duration T_C and b denotes the information symbol sent. The sequence of time-hopping gives an additional timeshift of $c_j T_C$ seconds to the jth pulse of the signal, in which c_j denote the Pseudorandom sequence elements, taking on integer values between 0 and $N_c - 1$. We perform matched filtering in the receiver, and sample this output at times $t = T_S + \tau_{run}$, and $t = T_S + \tau_{run} + T_d$. The sample values comparison at these two times computes which sequence had been transmitted.

In time hopping impulse radio, the suppression of noise is T_S/T_C.

Finally, the time hopping impulse radio is used not only in conjunction with pulse position modulation (PPM), but also with pulse amplitude modulation (PAM). Besides, the sent pulses polarity within a symbol can be done, so that the send signal then defines as follows –

$$s(t) = \frac{\sqrt{E_S}}{\sqrt{N_f}} \sum_{i} b_i \sum_{j=-\infty}^{\infty} d_j g(t - jT_f - c_j T_C - iT_S) \qquad \text{...(ii)}$$

where a Pseudorandom variable d_j multiplies with each pulse that can take on the values +1 or –1 with equal probability. Such a randomization of polarity has merits with respect to spectral shaping of the send signal. The resemblance to DS-CDMA is even more striking for such a system i.e., the spreading sequence is a ternary sequence now, with an equal number of +1's and –1's, and a larger number of 0's inbetween. Now, the interference suppression takes place not just due to a small number of pulse collisions, but also due to the interference contributions from distinct frames might cancel out.

Q.34. Explain the impulse radio in delay-dispersive channels in brief.

Ans. In coherent reception of the incoming signal, we require a Rake receiver. In essential case, the Rake receiver comprises multiple correlators or fingers. In each of the fingers, incoming signal correlation can be performed with a time-hopping sequence delayed version, in which the delay is proportional to the multi-path component delay that we want to "rake-up". Then, the Rake fingers output can be weighted and combined. Since impulse radio systems normally require large bandwidth, they always require structures in which fingers number is smaller than number of multipath components. For reasonable performance, it can be needed more Rake fingers even in indoor environments.

In transmitted reference (TR) schemes, the ideal matched filter in a channel of delay-dispersive can be matched to the time-hopping sequence convolution with the channel impulse response. We first have a filter matched to the time-hopping sequence for coherent reception, followed by Rake receiver, which is matched to the channel impulse response. This composite matched filter is produced by a transmitted reference scheme in a distinct manner. A transmitted reference transmitter sends out two pulses in each frame i.e., one unmodulated (reference) pulse, and a fixed time T_{pd} other, the modulated (data) pulse. The reference sequence pulse can be convolved with the impulse response of the channel when it is sent by the wireless channel. Hence, this signal constitutes a noisy version of the system impulse response and it should be noted that matched filtering is the same as correlating the received signal with this effective impulse response.

The receiver just multiplies the received signal with the received reference signal to perform a matched filtering on the received signal data carrying part. Now, the mathematical expression of the send signal for one symbol can be defined as –

$$p(t) = \sqrt{\frac{1}{2}}\sum_{j=0}^{N_f} d_j[g(t - jT_f - c_jT_C) + b.g(t - jT_f - c_jT_C - T_{pd})] \quad ...(i)$$

From equation (i), we observe that correlation with the received reference signal is performed by multiplying the complete received through a delayed version of itself and integration.

OO

UNIT 2 ORTHOGONAL FREQUENCY DIVISION MULTIPLEXING (OFDM)

INTRODUCTION, PRINCIPLE OF OFDM, IMPLEMENTATION OF TRANSCEIVERS, FREQUENCY-SELECTIVE CHANNELS, CHANNEL ESTIMATION, PEAK-TO-AVERAGE POWER RATIO

Q.1. Write an introductory note on orthogonal frequency division multiplexing (OFDM).

Ans. The orthogonal frequency division multiplexing is a wideband wireless digital communication scheme that is based on block modulation. Very high bit rate/high-speed communication is needed for video communication. The OFDM converts a high-rate datastream into a number of low-rate streams that can be transmitted over parallel, narrowband channels that are equalized easily.

In order to get the required data rate, the symbol duration T_S has to become very small, and the bandwidth of system becomes very large, when only a single band of frequency is used. Hence, the impulse response becomes very long in terms of symbol durations, when the symbol duration becomes very small. The probability of instabilities maximizes, and the computational effort can be very large for such a long equalizer.

Q.2. Discuss the working principle of OFDM. ***(R.G.P.V., June 2014)***

Or

Explain the principle of working of OFDM. ***(R.G.P.V., June 2012)***

Or

Discuss the spectrum of an OFDM signal. ***(R.G.P.V., Dec. 2011)***

Or

Discuss the principle of OFDM system. ***(R.G.P.V., June 2017)***

Ans. The concept of orthogonal signals is essential for the understanding of orthogonal frequency division multiplexing (OFDM) system. The information is divided into N parallel streams by the OFDM, which are then transmitted through modulating N distinct carriers. Hence, symbol duration becomes larger by a N factor on each subcarrier. Orthogonality is a property that permits multiple information carriers to be transmitted suitably over a common channel and detected without interference. A much narrower subcarriers spacing are achieved. Specifically, it is supposed that the subcarriers be at the frequencies $f_n = nW/N$, where W denotes the total available bandwidth and n denotes an integer. Besides, it is assumed that modulation on each of the subcarriers is PAM (pulse amplitude modulation) with rectangular pulse basis for the moment. Then, we see that subcarriers can be mutually orthogonal, because the relationship as follows –

$$\int_{iT_S}^{(i+1)T_S} e^{-j2\pi f_n t} e^{j2\pi f_k t} dt = \delta_{nk} \quad ...(i)$$

This principle of OFDM in the frequency domain is shown in fig. 2.1. The spectra of each modulated carrier has a sin (x)/x shape, because of the rectangular pulse shape in the time domain. The various spectra modulated carriers overlap, but each carrier is in the spectral nulls of all other carriers. Consequently, the datastreams of any two subcarriers will not interfere as long as the receiver does the suitable demodulation.

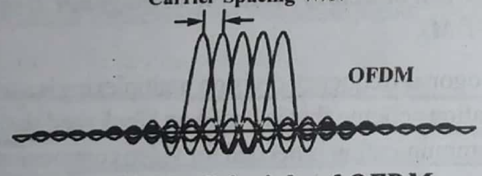
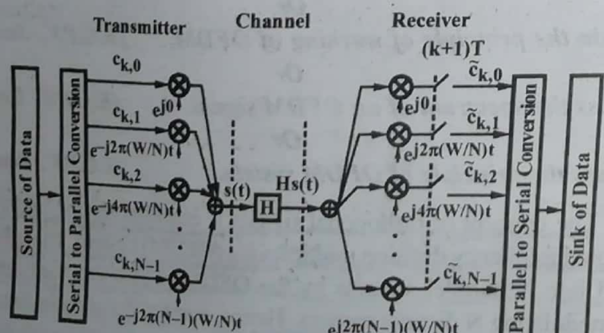


Fig. 2.1 The Principle of OFDM

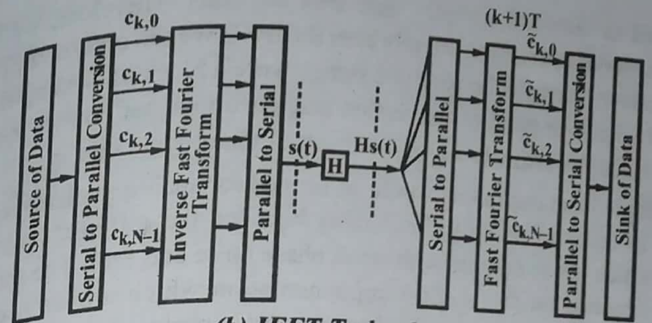
Q.3. *Explain the implementation of transceivers in OFDM.*
(R.G.P.V., June 2013)

Ans. There are two ways in which OFDM can be interpreted as follows –

(i) An ***analog interpretation*** is shown in fig. 2.2 (a). We first divide our original datastream into N parallel datastreams, each of which contain a lower rate of data. Besides, we have a number of local oscillators, at a frequency $f_n = nW/N$ each oscillates, where n = 0, 1, 2, ..., N – 1. Then, each of the datastreams modulates one of the carriers.



(a) Analog Interpretation



(b) IFFT Technology

Fig. 2.2 Transceiver Structure for OFDM

(ii) In an alternative implementation the transmit data splits into blocks of N symbols. This data block can be subjected to an inverse fast fourier transformation (IFFT), and then transmitted as shown in fig. 2.2 (b). With digital technology, this technique is very simple to implement.

These two techniques can be shown equivalent in following way –

The analog interpretation is considered first. Suppose, the $c_{n,i}$ is the complex transmit symbol at time instant **i** on the nth carrier then, the transmit signal can be as follows –

$$s(t) = \sum_{i=-\infty}^{\infty} s_i(t) = \sum_{i=-\infty}^{\infty} \sum_{n=0}^{N-1} c_{n,i}\, g_n(t - iT_S) \quad ...(i)$$

where, the basis pulse $g_n(t)$ is a normalized, frequency-shifted rectangular pulse as follows –

$$g_n(t) = \begin{cases} \dfrac{1}{\sqrt{T_S}} e^{\left(j2\pi n \frac{t}{T_S}\right)} & \text{for } 0 < t < T_S \\ 0 & \text{otherwise} \end{cases} \quad ...(ii)$$

Now, we can consider the signal only for i = 0 without restriction of generality and sample it at instances $t_k = kT_S/N$ as follows –

$$s_k = s(t_k) = \frac{1}{\sqrt{T_S}} \sum_{n=0}^{N-1} c_{n,0}\, e^{\left(j2\pi n \frac{k}{N}\right)} \quad ...(iii)$$

Consequently, an inverse discrete fourier transform (IDFT) is performed on the transmit block symbols to realize the transmitter. The IDFT can be realized as an IFFT, and the number of samples can be selected to be a power of 2 in all practical cases.

It should be noted that the IFFT input is made up of N samples, and for that reason the IFFT output also consists of N values. These N values are transmitted as temporal samples, one after the other. Therefore, we have a parallel-to-serial conversion directly after the IFFT. We can reverse the process at the receiver – sample the received signal, write a block of N samples into a vector i.e., serial-to-parallel conversion and perform a fast Fourier transform on this vector. An estimate $\tilde{c}_n$ is the result of the original data c_n.

Multiple local oscillators would be needed for analog implementation of OFDM, in order to retain orthogonality between the different subcarriers, each of which has to operate with small phase noise and drift. The success of OFDM depends on fast digital implementation which allows use of the transceivers which are much cheaper and simpler. In particular, the implementation of an FFT can be found by highly efficient structures, and only with log (N), the computational effort of performing an FFT maximizes.

Q.4. Draw the conceptual block diagram of OFDM systems and explain it in detail. ***(R.G.P.V., June 2015)***

Or

Explain in detail the block diagram of OFDM system and state the importance of each stage. ***(R.G.P.V., Dec. 2016)***

Ans. Refer to the ans. of Q.2 and Q.3.

Q.5. Explain OFDM. Give its types, characteristics features and its applications in communication systems. ***(R.G.P.V., June 2016)***

Ans. Refer to the ans. of Q.1, Q.2 and Q.3.

The types of OFDM are as follows –

(i) Flash OFDM (FOFDM)
(ii) Vector OFDM (VOFDM)
(iii) Wideband OFDM (WOFDM)
(iv) Adaptive OFDM (AOFDM)

(i) Flash OFDM – This type of OFDM uses multiple tones and fast hopping to spread signals over a given spectrum band.

(ii) Vector OFDM – It uses the concept of MIMO technology. MIMO uses multiple antennas to send and receive the signals so that multipath effects may be utilized to improve the signal reception and improve the transmission speeds which can be supported.

(iii) Wideband OFDM – It uses a degree of spacing between the channels which is large sufficient that any frequency errors between the sender and receiver do not affect the performance. WOFDM is specially applicable for Wi-Fi systems.

(iv) Adaptive OFDM – While communicating, the wireless channel effects on the OFDM received signal differ because of selection of different parameter values as well as existing signal-to-noise ratio condition at that time. Different number of subcarriers with same bandwidth, different allowable bandwidth with same number of subcarriers, various modulation mapping technique, different FFT points, variation in pilot power and pilot positions, variations in cyclic prefix interval etc. are considered as important OFDM parameters. The above parameters are made adaptive with SNR conditions of the channel. The SNR variations can increase because of distance variations between the sender and receiver and fast or slow mobility, exhibiting time varying channel.

Q.6. Discuss the OFDM demodulation. ***(R.G.P.V., Dec. 2011)***

Ans. Demodulation is performed to get back the required setting in the frequency domain. The FFT of each OFDM symbol is done to determine the original transmitted spectrum. Then, the phase angle of each transmission carrier is evaluated and converted back to the data word by demodulating the received phase (demapping). Then the data words are split back to the same pattern as the original bits to have the serial data again by parallel-to-serial conversion.

The following steps are used to demodulate the OFDM signal –

(i) Partition the input stream into vectors representing each symbol period.
(ii) Take the FFT of each symbol period vector.
(iii) Extract the carrier FFT bins and determine the phase of each.
(iv) Determine the phase difference for each carrier from one symbol period to the next.
(v) Decode each phase into binary data.
(vi) Sort the data into the appropriate order.

Q.7. Discuss how the OFDM operated in a frequency-selective channel.

Ans. Delay dispersion can lead to loss of orthogonality between the subcarriers, hence to ICI (inter carrier interference). It is fortunate that, a special type of guard interval eliminates both these negative effects, known as the *cyclic prefix (CP)*.

Cyclic Prefix – Firstly, we can define a new base function for transmission as follows –

$$g_n(t) = e^{j2\pi n\left(\frac{W}{N}\right)t} \quad \text{for } -T_{cp} < t < \hat{T}_S \qquad ...(i)$$

where, again W/N denotes the carrier spacing, and $\hat{T}_S = N/W$. Now, the symbol duration is $T_S = \hat{T}_S + T_{cp}$. According to the base function definition,

Scanned by CamScanner

the normal OFDM symbol can be transmitted for the duration $0 < t < \hat{T}_S$. A copy of the last part of the symbol can be transmitted during time $-T_{cp} < t < 0$ is shown in fig. 2.3. It means that $g_n(t) = g_n(t + N/W)$ – a fact that is straightforward verified through putting into equation (i). This prepended part of the signal is known as the ***cyclic prefix*** (CP).

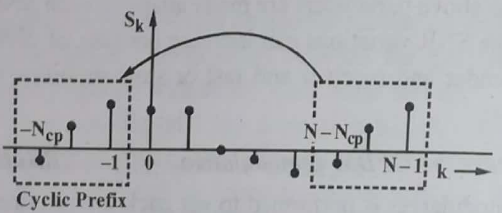


Fig. 2.3 Cyclic Prefix Principle

The arriving signal is the linear convolution of the transmitted signal with the impulse response of channel when any datastream is transmitted over a delay-dispersive channel. This linear convolution is converted into a cyclical convolution through the cyclic prefix. During the time $0 < t < \tau_{max}$, where τ_{max} denotes the maximum excess delay of the channel, because echoes of the last part of the preceding symbol interfere with the required symbol, the received signal subject to ***real*** ISI (inter symbol interference). During this time interval, this ISI can be removed through discarding the received signal. We have cyclical ISI during the remainder of the symbol. In particular, the current symbol last part that interferes with the current symbol first part.

It is assumed that the impulse response duration is equal to the cyclic prefix duration. Besides, we can assume i = 0 in order to make more simple notation. There is a filters bank that can be matched to the basis functions without the cyclic prefix as follows –

$$\bar{g}_n(t) = \begin{cases} g_n^*(\hat{T}_S - t) & \text{for } 0 < t < \hat{T}_S \\ 0 & \text{otherwise} \end{cases} \qquad \text{...(ii)}$$

The received signal's first part can be eliminated from the detection process by this operation. The matched filtering of the remainder is realized because an fast fourier transform operation. Hence, the signal is convolution of the transmit signal at the matched filter output with the receive filter and the impulse response of the channel as follows –

$$r_{n,0} = \int_0^{\hat{T}_S} \left[\int_0^{T_{cp}} h(t,\tau) \left(\sum_{k=0}^{N-1} c_{k,0}\, g_k(t-\tau) \right) d\tau \right] g_n^*(t)\,dt + n_n \qquad \text{...(iii)}$$

where, n_n denotes the noise at the matched filter output. It should be noted that the argument of g_k can reach values between $-T_{cp}$ and $\hat{T}_S$, that is the region of definition of equation (i). The $h(t, \tau) = h(\tau)$ when the channel is considered because constant during the time T_S, and we express as follows –

$$r_{n,0} = \sum_{k=0}^{N-1} c_{k,0} \int_0^{\hat{T}_S} \left[\int_0^{T_{cp}} h(\tau)(g_k(t-\tau))d\tau \right] g_n^*(t)\,dt + n_n \qquad \text{...(iv)}$$

The inner integral is expressed as follows –

$$e^{j2\pi t k \frac{W}{N}} \int_0^{T_{cp}} h(\tau)\, e^{-j2\pi\tau k \frac{W}{N}}\, d\tau = g_k(t)\, H\left(k\frac{W}{N}\right) \qquad \text{...(v)}$$

where, $H\left(k\frac{W}{N}\right)$ denotes the transfer function of the channel at the frequency kW/N. Besides, since during the time $0 < t < \hat{T}_S$, the basis function $g_n(t)$ are orthogonal as follows –

$$\int_0^{\hat{T}_S} g_k(t)\, g_n^*(t)\, dt = \delta_{kn}(t) \qquad \text{...(vi)}$$

The received signal samples r is expressed as follows –

$$r_{n,0} = H\left(n\frac{W}{N}\right) c_{n,0} + n_n \qquad \text{...(vii)}$$

Hence, a number of parallel non-dispersive fading channels represents the OFDM system, each with its own complex attenuation $H\left(n\frac{W}{N}\right)$. Hence, the system equalization becomes very easy – it just needs division at the frequency of subcarrier through the transfer function.

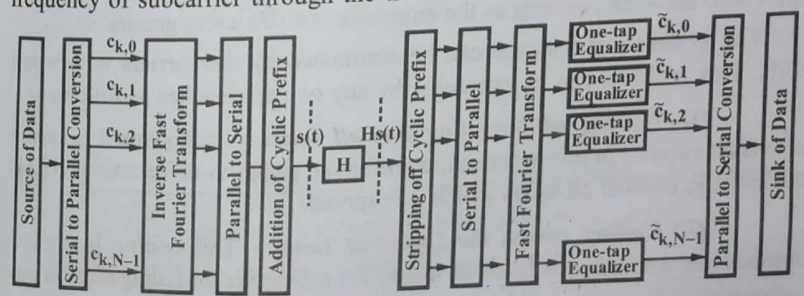


Fig. 2.4 OFDM Structure of Transmission Chain with One-tap Equalization and Cyclic Prefix

Fig. 2.4 shows the block diagram of an OFDM system including the cyclic prefix. The original datastream can be changed into S/P (serial to parallel). An IFFT subjects each block of N data symbols, and then the last NT_{cp}/T_S samples can be prepended. The resulting signal can be modulated onto a carrier and transmitted over a channel, which distorts the signal and adds noise. The signal can be divided into the blocks at the receiver. The cyclic prefix can be stripped off for each block, and the remainder can be experienced to an fast Fourier transform. The resulting samples can be ***equalized*** through one-tap equalization on each carrier i.e., division through the attenuation of complex channel.

Q.8. How time frequency selective channel is estimation is done in OFDM system ? **(R.G.P.V., May 2018)**

Ans. Refer to the ans. of Q.7.

Q.9. Explain the performance of OFDM in frequency selective channels.

Ans. A frequency-selective channel can be changed into a number of parallel flat-fading channels through the cyclic prefix. This is advantageous because it obtains rid of the ISI that plagues CDMA and TDMA systems. On the other hand the disadvantage is that any frequency diversity cannot be shown by an uncoded OFDM system at all. The probability of error on subcarrier is very high, when a subcarrier is in a fading dip, and for high signal-to-noise ratios, the BER (bit error rate) of the total system dominated.

We can easily achieved that uncoded OFDM contains the identical average BER regardless of the channel frequency selectivity. Also, this is interpreted as follows – variations of time provides us distinct realizations of channel at distinct times and frequency selectivity provides us distinct realizations of channel on distinct subcarriers. Doubly selective channels contain distinct realizations of channel on distinct subcarriers as well as at distinct times. But, it does not matter how the distinct realizations can be generated for computation of the average BER, as long as the ensemble is sufficiently greater.

The system performance can be dominated by the carriers with worst SNR. This problem can be overcome by any one approaches as follows –

(i) The Signal Spreading over all Tones – Each symbol is spread across all carriers in this approach, therefore an signal-to-noise ratio (SNR) is the average of over all tones which it is spread.

(ii) Coding across the Different Tones – This coding is used to compensate through a better SNR in another subcarrier for fading dips on one subcarrier.

(iii) Adaptive Modulation – When the transmitter understands the signal-to-noise ratio on each subcarriers it can select its adaptive modulation alphabet and adaptive coding rate. Hence, the transmitter will send using a smaller modulation alphabet and a stronger encoding on carriers with low signal to noise ratio. Also, the power allocated to each subcarrier is not content.

Q.10. What do you mean by coded orthogonal frequency division multiplexing in frequency selective channels ?

Ans. Sufficient interleaving should be applied so that fading of coded bits is not dependent. Alternatively, we require indepenent channel states over which to transmits our bits of coded. Automatically, this will result in a high order diversity. Channel states can be generated by either channel temporal variations, or because distinct transfer functions of subcarriers in a channels of frequency-selective. Hence, it is unnecessary to define new codes for OFDM, but it is a question on how to design suitable mappers and interleavers that assign the distinct coded bits in the plane of time-frequency. This mapping is dependent on the time selectivity as well as the frequency selectivity of the channel. It might be enough to code only across available frequencies, when the channel is highly frequency-selective, without any interleaving or coding along the time axis.

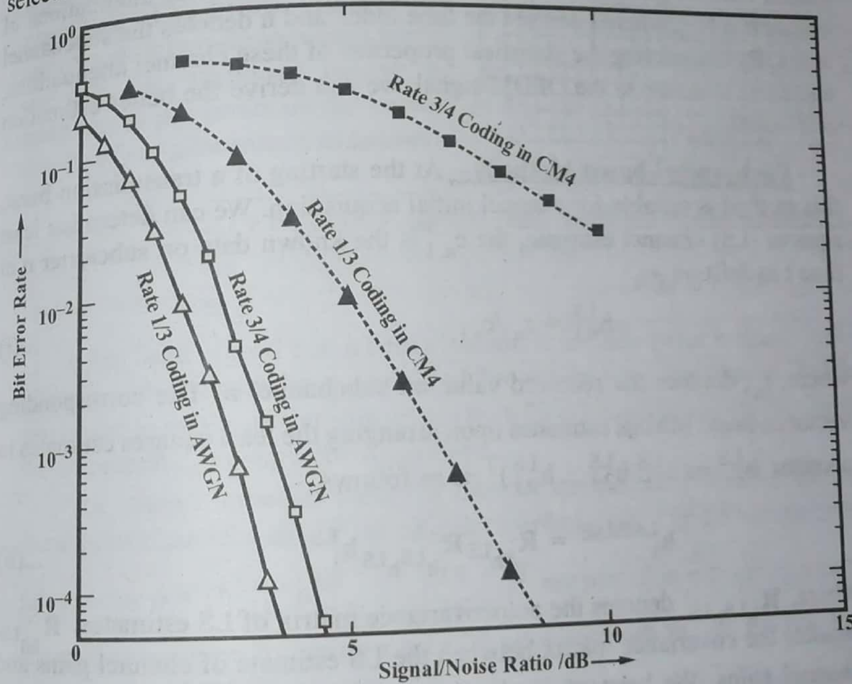


Fig. 2.5

A performance example is shown in fig. 2.5. We can see that, both rate-1/3 and 3/4 codes systems exhibit better performance, with approximately a 1 dB difference for AWGN. A performance is dramatically distinct in fading channels. While the code rate-1/3 contain better diversity, and consequently the bit error rate decrease fast as a SNR function, the code rate-3/4 contain very small frequency diversity, and hence the performance is worst.

Scanned by CamScanner

Advantages –

(i) The absence of interleaving results in lower latency of the transmission and decoding process in the time domain.

(ii) This technique also works in static channels, which happen quite often for wireless LANs and other high-rate data transmission situations.

Q.11. What do you understand by channel estimation in OFDM ?

(R.G.P.V., June 2013)

Ans. The OFDM systems operation needs an estimate the transfer function of the channel, or similarly, the impulse response of the channel, for any coherent wireless system. It is simple to estimate the channel in the frequency domain, because OFDM is operating with a number of parallel narrowband subcarriers. More accurately, we need to obtain estimates of the N complex-valued channel gains on the subcarriers. Suppose that, these attenuations of channels is $h_{n,i}$, where i denotes the time index and n denotes the subchannel index. By considering the statistical properties of these channel attenuations, and some structure to the OFDM signal, we can derive the better estimation of channel.

Pilot-symbol-based Methods – At the starting of a transmission burst, this method is suitable for channel initial acquisition. We can determine least squares (LS) channel estimate, the $c_{n,i}$ is the known data on subcarrier n at time t as follows –

$$h_{n,i}^{LS} = r_{n,i}/c_{n,i} \qquad ...(i)$$

where, $r_{n,i}$ denotes the received value on subchannel n. The corresponding vector of linear MMSE estimates upon arranging the least squares estimates in a vector $\mathbf{h}_i^{LS} = (h_{1,i}^{LS}\ h_{2,i}^{LS} \ldots h_{n,i}^{LS})^T$ is as follows –

$$\mathbf{h}_i^{LMMSE} = \mathbf{R}_{hh^{LS}} \mathbf{R}_{h^{LS}h^{LS}}^{-1} \mathbf{h}_i^{LS} \qquad ...(ii)$$

where, $\mathbf{R}_{h^{LS}h^{LS}}$ denotes the autocovariance matrix of LS estimates, $\mathbf{R}_{hh^{LS}}$ denotes the covariance matrix between the LS estimate of channel gains and channel gains. We have given that on each subcarrier with variance σ_n^2, $\mathbf{R}_{h^{LS}h^{LS}} = (\mathbf{R}_{hh} + \sigma^2\mathbf{I})$ and $\mathbf{R}_{hh^{LS}} = \mathbf{R}_{hh}$. Arranging the attenuations of channel in a vector $\mathbf{h}_i = (h_{1,i}\ h_{2,i} \ldots h_{n,i})^T$, we can obtain as follows –

$$\mathbf{R}_{hh} = E\{\mathbf{h}_i\mathbf{h}_i^{\dagger}\} = E\{\mathbf{h}_i^{*}\mathbf{h}_i^{T}\}^{*} \qquad ...(iii)$$

which does not depend on time i when the channel is wide-sense-stationary. It should be noted that, we have expressed $\mathbf{R}_{hh}$ because the conjugate of the "usual" correlation matrix in order to make more simple notation of the subsequent equations.

This estimation method gives very good results but become highly complex when the number of subcarriers is greater. It needs N^2 multiplications i.e., N multiplications per estimated channel gain. This method is quite complex, even if it is only applied at beginning of a transmission burst.

Q.12. Discuss the methods based on scattered pilots of channel estimation.

Ans. An attractive type of tracking the channel is to use pilot symbols scattered in the OFDM time-frequency grid as shown in fig. 2.6, where pilot can be spaced by N_t OFDM symbols and N_f subcarriers. We can begin through performing least squares estimation of the channel at pilot positions if estimating the channel depend upon the scattered pilots i.e., $h_{n,i}^{LS} = r_{n,i}/c_{n,i}$, where, $c_{n,i}$ denotes the known pilot data and $r_{n,i}$ denotes the received value in pilot position (n, i). We can perform interpolation to get channel estimates at all other positions from these initial estimates at pilot positions. The pilots interpreting samples in a two-dimensional space, we can use standard theory of sampling to put limits on the needed density of our pilot pattern as follows –

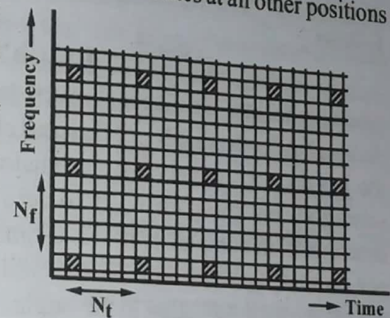


Fig. 2.6

$$N_t < \frac{1}{2(1+N_{cp}/N)\nu_{max}}$$

$$N_f < \frac{N}{N_{cp}}$$

It has been argued that a better tradeoff is to place twice because various pilots in each direction because desired through the sampling theorem, because we require to minimize the noise effect from the pilots and also help to minimize the complexity of estimation algorithms.

The same estimation theory is used in interpolating channels between these pilot channels, as for the all-pilot symbol case. The LS estimates in a pilot vector $\mathbf{p} = (h_{n_1,i_1}^{LS}\ h_{n_2,i_2}^{LS} \ldots h_{n_K,i_K}^{LS})^T$ are placed in estimating a certain channel attenuation $h_{n,i}$ using a set of K pilot positions (n_j, i_j), J = 1 ... K and compute the LMMSE estimate as –

$$h_{n,i}^{LMMSE} = \mathbf{r}_{hp}\mathbf{R}_{pp}^{-1}\mathbf{p}$$

where, $\mathbf{R}_{pp}$ denotes the $E\{\mathbf{pp}^{\dagger}\}$ and $\mathbf{r}_{hp}$ denotes the correlation vector $E\{h_{n,i}\ \mathbf{p}^{\dagger}\}$. This estimator complexity increases with the number of pilot tones comprised in the estimation and need K multiplication per estimated attenuation, again it is assumed that all correlation matrices and inversions can be determined before. A quite large number of pilots can be used to get better channel estimates.

Q.13. Discuss the methods based in eigen decompositions of channel estimations.

Ans. The matrix multiplication can be performed very efficiently by the channel statistical properties, if the LMMSE estimator is used in the following equation –

$$\mathbf{h}_i^{LMMSE} = \mathbf{R}_{hh^{LS}} \mathbf{R}^{-1}_{h^{LS}h^{LS}} \mathbf{h}_i^{LS} \quad ...(i)$$

From estimation theory, this is done with the theory of optimal rank minimization, where EVD (eigen value decomposition) $\mathbf{R}_{hh} = \mathbf{U}\Lambda\mathbf{U}^{\dagger}$ gives in a new efficient version of equation (i). The N_{cp}+1 is the dimension of this space i.e., the number of samples in the cyclic prefix is less than one. Consequently, the magnitude should be minimum rapidly, after the first N_{cp}+1 diagonal elements in Λ. Equation (i) can be written with the eigen value decomposition (EVD) as follows –

$$\mathbf{h}_i^{LMMSE} = \mathbf{U}\Delta\mathbf{U}^{\dagger}\mathbf{h}_i^{LS}$$

where, Δ denotes the diagonal matrix having the valves $\delta_i = \lambda_i/(\lambda_i + 1/\gamma)$ on its diagonal. After the first N_{cp}+1 diagonal elements δ_i will minimize rapidly because the λ_i's do. An optimal rank-p estimator for channel gains can be obtained by setting all but the p first λ_i's to zero i.e., assigning $\delta_i = 0$ for $i > p$. The determination complexity of this estimator is $2N_p$ multiplications, which is 2p per estimated attenuation. This should be compared with the multiplications per estimated attenuation in the original estimator equation (i). Fig. 2.7 shows the estimated principle, in which the optimal rank-p channel estimator seen as a transform (U^H) followed through p scalar multiplication and a second transform U.

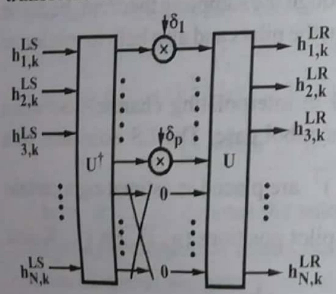


Fig. 2.7

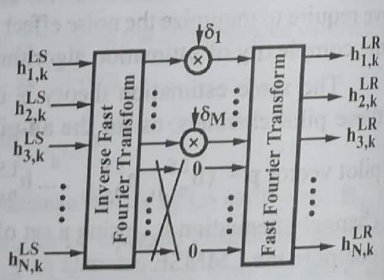


Fig. 2.8

There are only N_{cp} nonzero singular values, and the resulting optional transforms $\mathbf{U}^{\dagger}$ and $\mathbf{U}$ are the inverse discrete Fourier transform (IDFT) and discrete Fourier transform (DFT), in the case when the autocorrelation matrix $\mathbf{R}_{hh}$ is a circulant matrix. The basic estimator structure stays the similar, as shown in fig. 2.8, on the other hand the fast fourier transform in the OFDM is used to better perform channel estimation.

Generally, the computational efficiency of DFT-based estimators can greater the suboptimality of their rank minimization, when the channel correlation matrix is not circulant.

Q.14. Explain the origin of the peak-to-average-ratio problem. Also discuss its reduction techniques.

Ans. Peak-to-average ratio (PAR) issue begins from the fact that an OFDM signal is the superposition of N sinusoidal signals on distinct subcarriers. The emitted power is linearly proportional to N on average. Even though, the signals on the subcarriers add up constructively, so that the signal amplitude is proportional to N, and hence the power goes with N^2. Hence, we can anticipate the power PAR to maximize linearly with the subcarriers number.

Alternatively, the contributions to the total signal from the distinct subcarriers are seen as random variables. We invoke the central limit theorem to prove the distribution of the amplitudes of in-phase components is Gaussian, when the subcarrier, number is large, with a standard deviation $\sigma = 1/\sqrt{2}$ like unity mean power. The absolute amplitude is Rayleigh-distributed because both quadrature and in-phase components are Gaussian. Easily, compute the probability that the instantaneous amplitude will lie above a given threshold, knowing the amplitude distribution, and identically for power.

The amplitude distribution is important as it faces us with following choices are –

(i) The amplifier properties will cause to errors in the output signal with a nonlinear amplifier. Those nonlinear errors destroy orthogonality between subcarriers, and also cause to maximized out-of-band emissions.

(ii) Use a power amplifier into the transmitter that can increase linearly upto the possible peak value of the transmit signal. Because it needs expensive and power-consuming class-A amplifiers cannot be used practically. The subcarriers number N is larger, then, this solution become more difficult.

Peak-to-average Minimization Methods –

(i) Adjustments of Phase – The method first defines an ensemble of adjustment of phase vectors ϕ_l, $l = 1, 2, ..., L$, that can be known to both the transmitter and receiver. Each vector has N entries $\{\phi_n\}_l$. Then the transmitter multiplies the OFDM symbol to be transmitted c_n through each of these phase vector to obtain as follows – ...(i)

$$\{\hat{c}_n\}_l = c_n e^{j(\phi n)l}$$

and then chooses as follows – ...(ii)

$$\hat{l} = \arg\min_{l}(PAR(\{\hat{c}_n\}_l))$$

which provides the lowest PAR. Then, together with the index $\hat{i}$ transmits the vector $\{\hat{c}_n\}_{\hat{i}}$. Then, the receiver can undo the adjustment of phase, and demodulate the OFDM symbol.

(ii) Coding for PAR Minimization – Each OFDM symbol can represent one of 2^N codewords under normal conditions. Now, these of codewords only a subset of size 2^K is able to accept in the sense that its PAR is lower compare to given threshold. Both the transmitter and the receiver understand the mapping between a bit combination of length K, and the codeword length N that is selected to represent it, and which contain an admissible PAR. Hence the transmission method is as follows –

(a) The incoming bitstream parse into blocks of length K.

(b) The associated codeword of length N is selected.

(c) This codeword is transmitted through the OFDM modulator.

(iii) Correction through Multiplicative Function – The OFDM signal is multiplied by a time-dependent function whenever the peak value is large. The signal is multiplied by a factor A_0/S_k, when the signal reaches a level $s_k > A_0$. While, the transmit signal becomes as follows –

$$\hat{s}(t) = s(t)\left[1 - \sum_k \max\left(0, \frac{|s_k| - A_0}{|s_k|}\right)\right] \quad ...(iii)$$

when the level exceeds threshold, a less radical method is to multiply the signal through a Gaussian function centered at times, which is given below –

$$\hat{s}(t) = s(t)\left[1 - \sum_n \max\left(0, \frac{|s_k| - A_0}{|s_k|}\right) e^{-\frac{t^2}{2\sigma_t^2}}\right] \quad ...(iv)$$

Multiplication with a Gaussian function in the time domain with variance σ_t^2 means convolution by a Gaussian function of variance $\sigma_f^2 = 1/(2\pi\sigma_t^2)$ in the frequency domain. Hence, the judicious choice of σ_t^2 influences the amount of out-of-band interference.

(iv) Correction by Additive Function – We can select an additive, similarly, rather a multiplicative, correction function. The correction function should be smooth sufficient not to include significant out-of-band interference. Besides, the correction function actions as additional pseudo-noise, and hence the BER of the system is maximized.

Q.15. Derive a relation for peak to average power ratio.

(R.G.P.V., June 2016)

Ans. The peak power and an average power are required to derive an expression for peak-to-average power ratio in OFDM. In order to calculate the largest possible peak power of an OFDM signal x(t), we consider without loss of generality that the N subcarriers are modulated with a random sequence of the real coefficients $c_{ik} = \pm 1$. The average power of such a random OFDM signal is equal to the sum of the average powers of the N orthogonal subcarriers. The average power for $c_{ik} = \pm 1$ has to be divided by a format dependent factor k^2 for arbitrary modulation coefficients c_{ik}. Then the peak-to-average power ratio (PAPR) in OFDM is given by

$$\boxed{PAPR_{OFDM} = \frac{N^2}{\frac{1}{2}N/k^2} = 2\,Nk^2} \quad ...(i)$$

With increasing number N of subcarriers the value for $PAPR_{OFDM}$ increases linearly.

For a finite approximation of a sinc-impulse only R_O/q pulses may contribute. Here R_O represents the number of time intervals T_s/q for q-fold oversampling i.e., R_O stands for the length of the impulse response. We determine the maximum power with $Q = R_O/q$ as –

$$P_{max} = \left[\sum_{-R/(2q)+1}^{R/(2q)} \left|\text{sinc}\left(\frac{1}{2} - r\right)\right|\right]^2 \quad ...(ii)$$

For large filter orders R_O, P_{max} could become arbitrarily large. Yet, while P_{max} increases with R_O, the probability for determining R_O sinc-pulses interfering constructively decreases as well similarly to the OFDM case. The average power $\overline{P}$ of an ideal Nyquist signal is given by

$$\overline{P} = \frac{1}{T_s}\int_{-\infty}^{+\infty} \text{sinc}\left(\frac{t}{T_s}\right) dt = 1 \quad ...(iii)$$

When Nyquist pulses are produced with a filter of finite order R, orthogonality as implied by equation (iii) is still a good assumption, then the average power of truncated Nyquist sinc-impulses is close to $\overline{P} = 1$. Equation (iii) has to be divided by a format dependent factor k^2 for arbitrary modulation coefficients c_{ik}. The PAPR for Nyquist signal is defined as the ratio of maxim power to average power, i.e.,

$$\boxed{PAPR_{Nyquist} = \frac{P_{max}}{\overline{P}}}$$

$$\simeq k^2 \left[\sum_{-R/(2q)+1}^{R/(2q)} \left| \text{sinc}\left(\frac{1}{2} - r\right) \right| \right]^2 \quad ...(i$$

Unlike OFDM signals in which the PAPR rises linearly, the PAPR of Nyquist signals does not, due to the temporal decay of its elementary sinc impulse.

Q.16. How is channel estimation done in OFDM ? Determine the peak-to-average power ratio. ***(R.G.P.V., June 2014)***

Or

Discuss channel estimation in OFDM. Also determine the peak to average power ratio. ***(R.G.P.V., June 2015)***

Ans. Refer to the ans. of Q.11 and Q.15.

Q.17. What is input backoff ? ***(R.G.P.V., Dec. 2016)***

Ans. The input backoff (IBO) can be defined as

$$\boxed{IBO = 10 \log_{10} \frac{P_{in}^{max}}{P_{in}}}$$

where P_{in} = Input power

P_{in}^{max} = Maximum possible input power.

INTERCARRIER INTERFERENCE, ADAPTIVE MODULATION AND CAPACITY, MULTIPLE ACCESS, MULTICARRIER CODE DIVISION MULTIPLE ACCESS, SINGLE CARRIER MODULATION WITH FREQUENCY-DOMAIN EQUALIZATION

Q.18. What is intercarrier interference ? How can it be reduced ? ***(R.G.P.V., June 2014)***

Or

What is intercarrier interference ? Explain how it can be overcome. ***(R.G.P.V., June 2016)***

Or

What is intercarrier interference and how it can be overcomed ? ***(R.G.P.V., June 2017)***

Ans. Wireless propagation channels are time-varying, and hence time-selective. Time selectivity contains two important results for an OFDM are as follows –

(i) It generates the intercarrier interference (ICI).

(ii) It causes to random frequency modulation, which can lead distortions on subcarriers that are in a fading dip.

Fig. 2.9 shows a Doppler shift of one subcarrier can lead ICI in various adjacent subcarriers. The product of maximum symbol duration and Doppler frequency of the OFDM computes the effect of time selectivity. The symbol duration is inversely proportional to the spacing between the subcarriers. Hence, when symbol duration is large, even a small Doppler shift can give in large intercarrier interference (ICI).

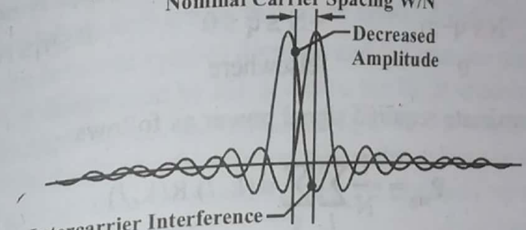


Fig. 2.9 Intercarrier Interference Because of Frequency Offset

When the cyclic prefix is smaller compare to the maximum excess delay of the channel, then the delay dispersion is another source of intercarrier interference (ICI). In order to improve spectral efficiency, a system might aware of shorter or leave out the cyclic prefix. In other cases, a system can be designed to work in a specific class of environments is later also deployed in greater excess delay. Finally, a compromise between the require to retain spectral efficiency and the require to avoid ICI is the length of the cyclic prefix for various systems. On the other hands, a cyclic prefix cannot be selected with the bad case channel condition.

The intercarrier interference methods can be classified as follows –

(i) ***Optimum Choice of OFDM Basis Signal*** – In order to decrease intercarrier interference (ICI), a method influences the OFDM basis pulse shape. A rectangular temporal signal shows a very sharp cut-off in the temporal domain, but shows a sin(x)/x shape in the frequency domain, and hence slowly decays. The slope of the sin (x)/x is large near its zeros, and on the other hand in a perfect system each subcarrier is in the spectral nulls of all other subcarriers. Hence, even a small Doppler shift causes to more intercarrier interference (ICI). The intercarrier interference (ICI) due to the Doppler shift can be reduced by selecting basis pulse whose spectrum decays higher and gentler. On the

downside, faster decay in the frequency domain can be bought through sl decay in the time domain, which maximize delay-spread induced errors.

(ii) Optimum Choice of Carrier Spacing and OFDM Sym Length – In order to decrease intercarrier interference (ICI), a related me influences the OFDM symbol length. As stated above that short symbol dur is better for minimization of Doppler-induced intercarrier interference. Wh spectral efficiency considerations enforce a decrease duration of T_S. T cyclic prefix should be greater than 10% of the symbol duration. The samp delay cross power spectral density is the $R(k, l) = P_h(lT_c, kT_c)$. Besides function can be expressed as follows –

$$w(q, r) = \frac{1}{N}\begin{cases} N-|r| & 0 \le q \le N_{cp} & 0 \le |r| \le N \\ N-q+N_{cp}-|r| & N_{cp} \le q \le N+N_{cp} & 0 \le |r| \le N-q+N_{cp} \\ N+q-|r| & -N \le q \le 0 & 0 \le |r| \le N+q \\ 0 & \text{elsewhere} & \end{cases} \quad \ldots(i)$$

Then, approximate required signal power as follows –

$$P_{sig} = \frac{1}{N}\sum_{l}\sum_{k} w(k, l)\, R(k, l) \quad \ldots(ii)$$

and the intercarrier symbol interference (ISI) and intercarrier interference (ICI) powers as follows –

$$P_{ISI} = \sum_{l}[1 - w(k, 0)]\, R(k, 0) \quad \ldots(iii)$$

$$P_{ICI} = \sum_{k} w(k, 0)\, R(k, 0) - P_{sig} \quad \ldots(iv)$$

and the signal-to-interference-and-noise-ratio (SINR) can be defined as follows-

$$SINR = \frac{\frac{E_S}{N_0} P_{sig} \frac{N}{N_{cp}+N}}{\frac{E_S}{N_0} P_{sig} \frac{N}{N_{cp}+N} \frac{P_{ISI}+P_{ICI}}{P_{sig}} + 1} \quad \ldots(v)$$

The above equations permit a simple tradeoff between the ICI because of residual delay dispersion, the ICI because of the Doppler effect and SNR loss because of the cyclic prefix.

(iii) Self-interference Cancellation Methods – A related method modulates the information not just onto a single subcarrier but onto a group of them. This method is more effective for mitigation of intercarrier interference (ICI), but causes to a minimization in system spectral efficiency.

Q.19. What is intersymbol interference ? **(R.G.P.V., Dec. 2016)**

Ans. For different MPCs, the runtimes are different. In narrowband systems, this can lead of different phases of MPCs, which leads to interference. A system with large bandwidth has good resolution in time domain. The signal dispersion is the major result. It means that the impulse response of the channel is not a single delta pulse, but a chain of pulses. Each of the pulse has a different arrival time, and a different amplitude and phase as well as shown in fig. 2.10. At the receiver, this signal dispersion leads to intersymbol interference. The kth bit information is carried by MPCs with long runtimes, and contributions from (k + 1)th bit are carried by MPCs with short runtimes arrive at the same time at the receiver, and interfere with each other as shown in fig. 2.11. Assuming that, this inter symbol interference leads to errors that cannot be eliminated by simply increasing the transmit power, if no special measures are taken, and are so that often known irreducible errors.

The ratio between symbol duration and the impulse response duration of the channel is determined by ISI. It means the ISI is important for higher data rates, and multiple access methods leading to an increase in transmitted ***peak data*** rate. Finally, it should also be noted that when the impulse response duration is shorter than bit duration, then ISI can also play a role.

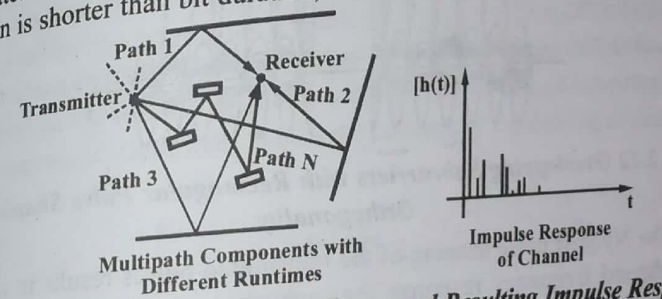


Fig. 2.10 Propagation of Multipath Waves and Resulting Impulse Response

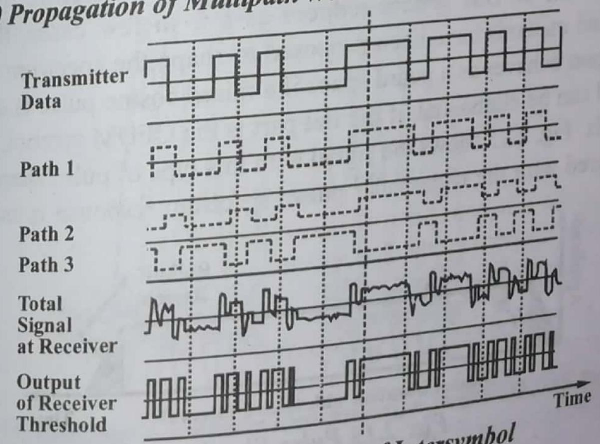


Fig. 2.11 Interference of Intersymbol

Q.20. Discuss pulse shaping, windowing and synchronization issues in OFDM. **(R.G.P.V., June 2013)**

Ans. Pulse Shaping in OFDM Signal and Spectral Efficiency – Each individual subcarrier channel width as well as overall transmission bandwidth both must be shaped so that out-of-band components can be reduced for improving the spectrum efficiency. At input stages, Nyquist pulse shaping is applied to input signal for narrow subcarrier channels. Final OFDM spectrum setting is obtained thereafter as shown in fig. 2.12.

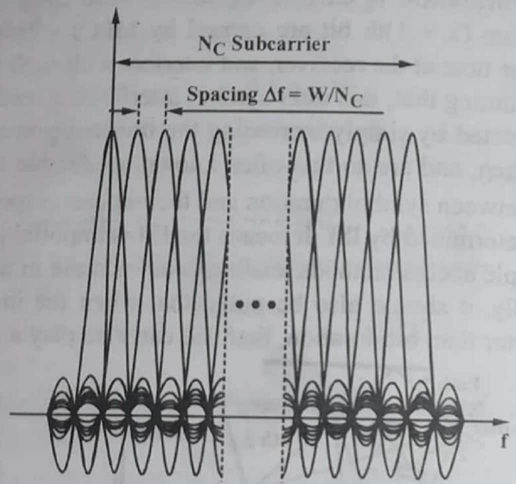


Fig. 2.12 Overlapping Subcarriers with Rectangular Pulse Shaping and Orthogonality

The Nyquist pulse shaping of the transmitted pulses results in a desired sinc-shaped frequency response for each channel. Therefore, the power spectrum of the OFDM system reduces as f^{-2}. In few cases, this is not adequate and methods have been proposed to shape the spectrum. Also, the roll-off region behaves as a guard space if a raised cosine pulse is used. Both ICI and ISI can be eliminated, if the flat part is the OFDM symbol, including cyclic prefix. Fig. 2.13 shows the signal with this type of pulse shaping, when it is compared with the rectangular pulse, spectrum response is as shown in

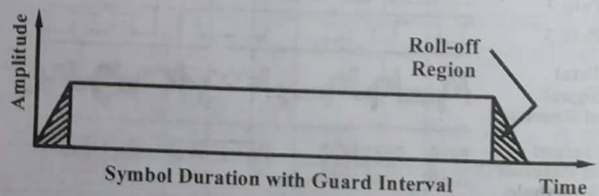


Fig. 2.13 Pulse Shaping

fig. 2.14. Since the spectrum falls much more quickly and decreases the interference to adjacent frequency bands, the overhead introduced by an extra guard band with a grace ful roll-off can be a good investment for ISI elimination.

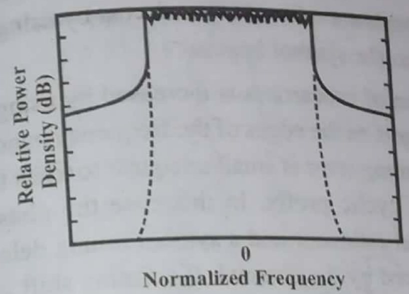


Fig. 2.14 Spectrum Shape

Windowing in OFDM Signal and Spectral Efficiency – If the overall spectrum of OFDM is concerned, due to the sinc shape of each narrowband channel, the side lobes with main lobe will appear. These are unwanted components and are to be removed from the final transmission bandwidth. This is obtained by applying the windowing function to final OFDM baseband signal in which effect of all the subcarriers is available. Passing OFDM baseband signal through the windowing block will perform the task of removing out of band components. Dark lines shows the concept of windowing as shown in fig. 2.15.

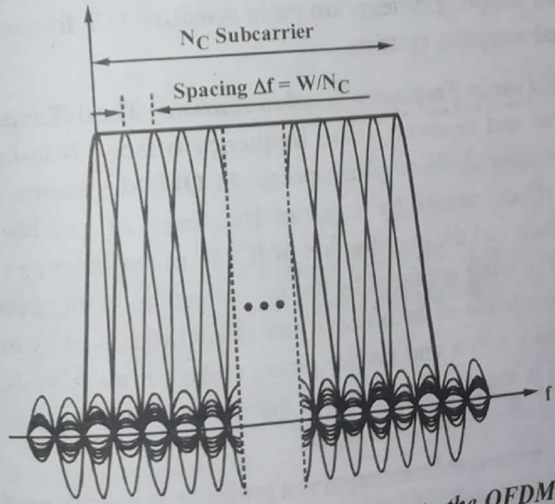


Fig. 2.15 Applying the Windowing Function to the OFDM Baseband Signal

Synchronization in OFDM Signal – Three synchronization problems occur –

(i) Timing Errors and Symbol Synchronization – A great deal of attention is provided to symbol synchronization in OFDM system. Although the timing requirements are relaxed to somewhat by using a cyclic prefix. The aim is to know when the symbol begins.

A phase rotation of subcarriers is increased by using timing offset. This phase rotation is largest on the edges of the frequency band. The orthogonality is maintained, if a timing error is small adequate to keep the channel impulse response within the cyclic prefix. In this case the phase rotations can be activated by a channel estimator and a symbol timing delay can be viewed as a phase shift introduced by the channel. If a timing shift is larger compare to the cyclic prefix, then ISI will take place.

(ii) Sampling Frequency Synchronization – The received continuous-time signal is sampled at the instants obtained by the receiver clock. Methods of dealing with the mismatch in sampling frequency are of two types –

(a) In synchronized-sampling system, a timing algorithm controls a VCO in order to align the receiver clock with the transmitter clock.

(b) In non-synchronized sampling, where the sampling rate remains fixed, which requires post processing in the digital domain.

The effect of the clock frequency offset is twofold-the useful signal component is rotated and attenuated and moreover ICI is introduced. Non-synchronized sampling systems are more sensitive to a frequency offset than synchronized sampling systems.

(iii) Carrier Frequency Synchronization – The differences in oscillator in transmitter and receiver create frequency offsets. Non-linear channels introduce Doppler shifts or phase noise. In OFDM systems, there are two destructive effects caused by a carrier frequency offset. The reduction of signal amplitude and the introduction of ICI from the other carriers as shown in fig. 2.16. The latter is caused by the loss of orthogonality between the sub-channels. Degradation of the bit error rate (BER) is caused by the presence of carrier frequency offset and carrier phase noise for an AWGN. Thus, multi carrier system is much more sensitive compare to the single-carrier system. The frequency offset is denoted by δf.

If the frequency synchronization is a problem, it is decreased by lowering the number of subcarriers, which will increase the subcarrier spacing. Although, this will increase the demands on the time synchronization, since the symbol length get shorter, i.e., a larger relative timing error will take place. Hence, the synchronization between time and frequency are closely related to each other.

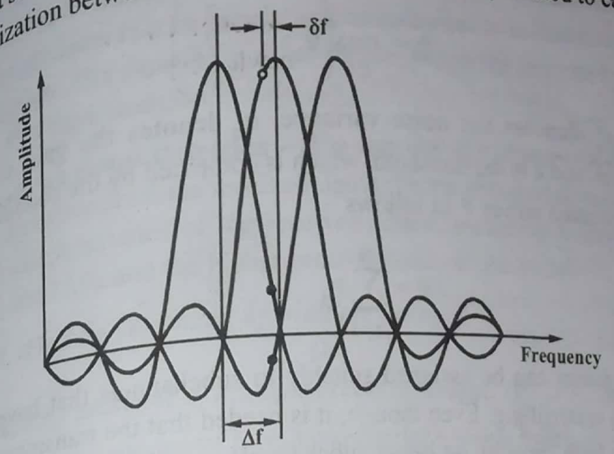


Fig. 2.16 Effect of Frequency Offset δf

Q.21. Write short note on adaptive modulation and capacity.
(R.G.P.V., June 2017)

Or

What is adaptive modulation and how does it affect the capacity ?
(R.G.P.V., June 2012, 2015)

Ans. The modulation/coding method can be changed by adaptive modulation based on channel – state information. For each subcarrier selects the modulation/coding distinctly in OFDM, and also it can vary with time. Obviously, the system can be performed better with adaptive modulation compare to the system whose coding and modulation can be fixed once and for all.

According to the channel quality, a code rate and modulation scheme can be chosen by adaptive modulation of a certain subcarrier.

(i) The Estimation of Channel Quality – It is needed adaptive modulation that the transmitter should knows channel-state information. It is needed that the channel to be reciprocal. While base station (BS) is in receiving mode, acquires knowledge of the channel state. It depends on the fact that the channel is still in the similar state, when it then transmits. This is only achieved in system with time domain duplex (TDD) in slowly time varying channels.

(ii) Adaptation of Parameter – The correct parameter for each subcarrier such as modulation alphabet and coding rate are decided by the transmitter after the channel state known. Besides, power allocation P_n of the nth subchannel was discovered by Shannon in the 1940s, and is known as ***waterfilling***. Which is as follows –

$$P_n = \max\left(0, \varepsilon - \frac{\sigma_n^2}{|\alpha_n|^2}\right) \quad ...(i)$$

where, σ_n^2 denotes the noise variance, α_n denotes the gain of the nth subchannel, and ε is the threshold, which is computed by the condition of the total transmitted power P as follows –

$$P = \sum_{n=1}^{N} P_n \quad ...(ii)$$

The power can be assigned suitably to subchannels that have a better SNR with waterfilling. Even though, it is needed that the transmitter, should use of the high capacity on better subchannels.

Signaling as close to capacity as possible should be done in each subchannel. It implies that the transmitter has to adapt the data rate according to the available SNR. The $\log_2 (N_a)$ limits the capacity per subchannel, where N_a denotes the symbol alphabet size. Hence, it is wasteful to allocate more energy to one stream than can be exploited by the alphabet actually. A ***providing to the poor*** principle for power allocation can be suitable, when the available alphabet is small.

(iii) Signaling of Selected Parameters – Practically, the transmitter has a finite and discrete set of modulation alphabets available. Also, there is a finite set of possible code rates. The distinct codes can be found from the ***mother*** code through distinct puncturing amounts.

The transmitter decides about the mode of transmission, i.e., a combination of signal encoder and constellation to use on each tone, and then communicates the decision to the receiver. Three techniques to get that task are as follows–

(a) Explicit Transmission – The transmitter can send the index of the transmission mode it intends to use in a robust and predefined format. This information transmission itself is performed in the same mode, and during transmission care should be taken that the message is well-protected against errors.

(b) Implicit Transmission – When the transmitter determines its channel-state information from the receiver by feedback, then implicit transmission is possible. In this case, what channel-state information is available to the transmitter can be known by the receiver, and this is the basis on which the decision can be made for a transmission mode. Hence, the receiver requires to know the decision rule on which the transmitter bases its mode of selection of transmission. The situation is more simpler, when the receiver feeds back the mode that the transmitter should use.

(c) Blind Detection – The receiver can attempt to obtain the signal constellation from the received signal. This is obtained through distinct statistical characteristics of the received signal, including autocorrelation functions, the PAR, and the higher order statistics of the signal.

Q.22. What do you mean by multiple access ?

Or

Explain the principle of multiple access technique. (R.G.P.V., June 2016)

Ans. The transmission of high data rates for a single user is permitted by the OFDM. The OFDM is a modulation format, and it cannot be viewed as a multiaccess format. Although, it can be used for assigning distinct subcarriers to distinct users. Especially, a number of subcarriers which are not adjacent to each other could be assigned to each user and hence offer sufficient frequency diversity. However, such a strategy leads to following problems –

(i) Signal from distinct users have distinct runtimes in the uplink, and thus do not reach at the receiver in a synchronized manner.

(ii) The administrative effort is high for such an assignment.

(iii) Each users ***sees*** distinct channels, destroying orthogonality between the signals from distinct users.

Instead, it is usual to combine OFDM with TDMA (time division multiple access) as the modulation formats. In that manner, full-frequency diversity can be obtained, on the other hand we can still retain orthogonality between the subcarriers. Alternatively, we can combine OFDM with FDMA (frequency division multiple access), where we can allocate each user a group of adjacent tones, separated through the spaces of frequency guard.

Q.23. Discuss the principle of multicarrier code division multiple access.
(R.G.P.V., June 2014, 2017)

Or

Explain what is multicarrier code division multiple access.
(R.G.P.V., June 2012)

Or

Describe multicarrier code division multiple access in detail.
(R.G.P.V., May 2018)

Ans. The information from each data symbol over all tones of an OFDM symbol is spread by multicarrier CDMA (MC-CDMA). It is absurd to combine multicarrier schemes, which attempts to signal over a narrower channel, with CDMA which attempt to spread a signal over a large bandwidth.

The MC-CDMA transmits a data symbol simultaneously on all available subcarriers. Hence, a code symbol c can be mapped to a vector **cp**, where **p** denotes a predetermined vector. This is interpreted as a repetition code or as a spreading action, where a code sequence represents each symbol. We obtain a bandwidth expansion through a factor of N regardless of the interpretation.

Bandwidth expansion causes a loss of spectral efficiency. Although, this problem can be removed by transmitting N distinct symbols, and hence N distinct codevectors, simultaneously. The first symbol c_1 is multiplied with codevector $\mathbf{p}_1$, the second symbol c_2 is multiplied with codevector $\mathbf{p}_2$, the third symbol c_3 is multiplied with codevector $\mathbf{p}_3$, and so on. The receiver can recover the distinct transmitted symbols, when all the vectors $\mathbf{p}_n$ are selected to be orthogonal.

Now, the codevectors can be written into a ***spreading matrix*** **P** as follows–

$$\mathbf{P} = [\mathbf{p}_1 \quad \mathbf{p}_2 \quad \mathbf{p}_3 \; \cdots \; \mathbf{p}_N] \qquad ...(i)$$

A matrix multiplication can be done by the ***symbol spreader*** as follows–

$$\tilde{\mathbf{c}} = \mathbf{Pc} \qquad ...(ii)$$

Now, this modified signal is OFDM modulated. It goes through an IFFT (inverse fast Fourier transform) and has the cyclic prefix prepended, and sent over the wireless channel.

In the receiver we can do the operation of stripping off the cyclic prefix and performing an FFT. Hence, the symbols again have been transformed into frequency domain. The received symbols at this stage are as follows –

$$\tilde{\mathbf{r}} = \mathbf{H}\tilde{\mathbf{c}} + \mathbf{n} \qquad ...(iii)$$

where **H** denotes the diagonal matrix with entries $H\left(n\frac{W}{N}\right)$ along the diagonal.

One-tap equalization is the second step. It is assumed that we use zero-forcing equalization for the moment. Then, we can perform despreading by multiplying the received signal with the Hermitian transpose of the spreading matrix by using the unitary properties of the spreading matrix to achieve as follows –

$$\mathbf{P}^{\dagger}\mathbf{H}^{-1}\tilde{\mathbf{r}} = \mathbf{P}^{\dagger}\mathbf{H}^{-1}\mathbf{HPc} + \mathbf{P}^{\dagger}\mathbf{H}^{-1}\mathbf{n} \qquad ...(iv)$$

$$= \mathbf{c} + \tilde{\mathbf{n}} \qquad ...(v)$$

Hence, the transmit symbols are recovered. Fig. 2.17 shows the block diagram of a multicarrier CDMA transceiver.

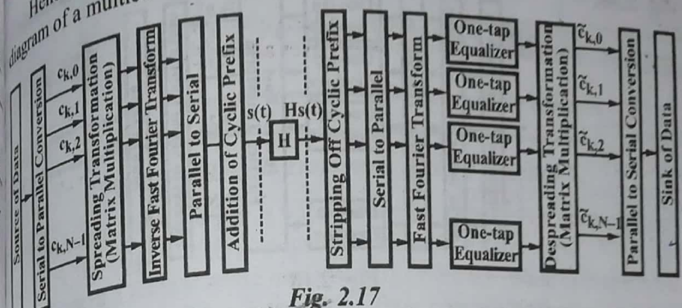


Fig. 2.17

It should be noted that, because the $\tilde{\mathbf{n}}$ comprises noise enhancement from zero-forcing equalization, noise is no longer white. The noise enhancement is good, if the receiver uses MMSE (minimum mean sequence error) equalization rather zero-forcing. Even though, MMSE equalization does not recover the orthogonality of distinct codewords like zero-forcing does. After equalization and despreading we will obtain –

$$\mathbf{P}^{\dagger}\frac{\mathbf{H}^*}{|\mathbf{H}|^2+\sigma_n^2}\tilde{\mathbf{r}} = \mathbf{P}^{\dagger}\frac{\mathbf{H}^*\mathbf{H}}{|\mathbf{H}|^2+\sigma_n^2}\mathbf{Pc} + \mathbf{P}^{\dagger}\frac{\mathbf{H}^*}{|\mathbf{H}|^2+\sigma_n^2}\mathbf{n} \qquad ...(vi)$$

The matrix $\frac{\mathbf{P}^{\dagger}\mathbf{H}^*\mathbf{H}}{(|\mathbf{H}|^2+\sigma_n^2)\mathbf{P}}$ is not diagonal and there is residual crosstalk from one codeword to the other codeword.

Q.24. Write short note on single carrier modulation with frequency domain equalization. ***(R.G.P.V., June 2017)***

Or

Explain the principle of single carrier modulation with frequency domain equalization. ***(R.G.P.V., June 2016, 2017)***

Or

What is the importance of single carrier modulation with frequency domain equalization ? ***(R.G.P.V., June 2015)***

Or

What is single carrier modulation with frequency domain equalization ? ***(R.G.P.V., May 2018)***

Ans. When FFT matrix is selected to be a unitary transformation matrix, a special case of MC-CDMA occurs in which multiplication through the spreading matrix and the inverse FFT inherent in the implementation of OFDM cancel out each other. While, the transmit sequence that can be transmitted over the channel is the original data sequence with a cyclic prefix. A block diagram of a single carrier frequency domain equalization receiver is shown in fig. 2.18.

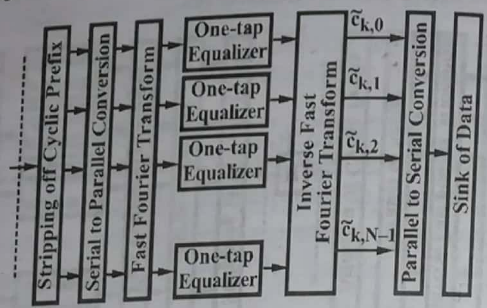


Fig. 2.18

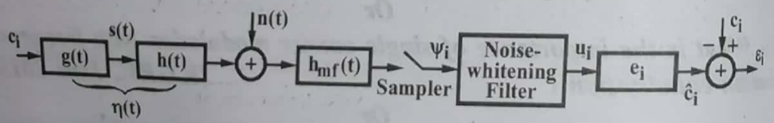
At the receiver, the signal is transformed by an FFT into the frequency domain after stripping off the cyclic prefix. There are no residual effects of ICI (inter carrier interference) or ISI (inter symbol interference) because of the cyclic prefix. Then, the receiver works equalization on each subcarrier, and finally an IFFT transforms the signal back into the time domain. Hence, the receiver performs equalization in the frequency domain. The computational effort for equalization goes only such that $\log_2(N)$, since IFFTs or FFTs is implemented efficiently. It is disadvantageous that the frequency domain equalization is a linear equalization technique and consequently does not provide designed performance.

Q.25. *What do you mean by linear equalization ? (R.G.P.V., Dec. 2016)*

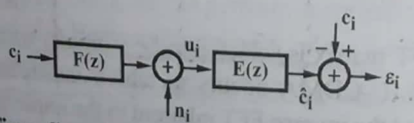
Ans. These are simple linear filter structures that attempt to invert the channel in the sense that the product of the transfer functions of channel and equalizer meets a some specific criterion.

This criterion is obtained either by minimizing the mean-squared error at the filter output or by a completely flat transfer function of the channel-filter concatenation.

Fig. 2.19 shows the basic structure of a linear equalizer.



(a) Structure of a Linear Equalizer in the Time Domain



(b) Structure of a Time-discrete Equivalent System in the Z-transform Domain

Fig. 2.19

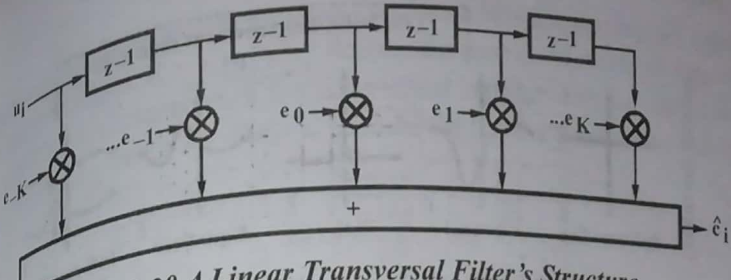


Fig. 2.20 A Linear Transversal Filter's Structure

A transmit sequence $\{c_i\}$ is transmitted over a dispersive, noisy channel, therefore the sequence $\{u_i\}$ is present at the equalizer input. Now, the coefficients of a finite impulse response (FIR) filter are obtained (transversal filter, fig. 2.20) with 2K + 1 taps. This filter converts sequence $\{u_i\}$ into sequence $\{\hat{c}_i\}$ a follows –

$$\hat{c}_i = \sum_{n=-k}^{k} e_n u_{i-n} \qquad \text{...(i)}$$

that is as close as possible to the sequence $\{c_i\}$.

The deviation ε_i is defined as follows –

$$\varepsilon_i = c_i - \hat{c}_i \qquad \text{...(ii)}$$

Therefore, we try to obtain a filter as follows –

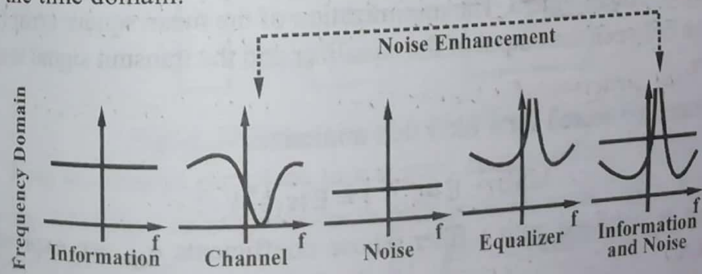
$$\varepsilon_i = 0 \text{ for } N_0 = 0 \qquad \text{...(iii)}$$

which provides the zero-forcing equalizer, or that –

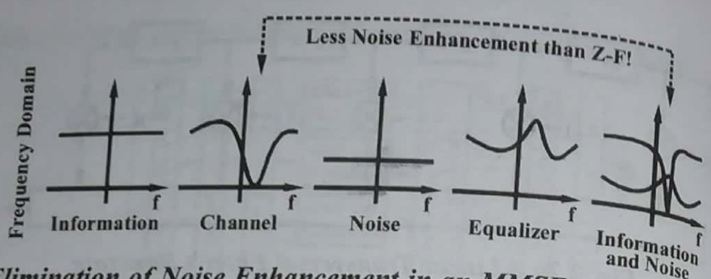
$$E\{|\varepsilon_i|^2\} \rightarrow \min \text{ for } N_0 \text{ having a finite value} \qquad \text{...(iv)}$$

which is providing the minimum mean square error (MMSE) equalizer.

(i) Zero-forcing Equalizer – The ZF equalizer is interpreted in the frequency domain because applying a completely flat (constant) transfer function of the combination of equalizer and channel by selecting the equalizer transfer function as $E(z) = 1/F(z)$. This is interpreted as minimizing the maximum ISI in the time domain.



(a) Noise Enhancement in Zero-forcing Equalizer



(b) Elimination of Noise Enhancement in an MMSE Linear Equalizer

Fig. 2.21

For reducing inter symbol interference, the zero-forcing equalizer is optimum. Even though, channels also add noise, which is amplified using the equalizer. The equalizer contains a larger amplification at frequencies where the channel transfer function receives low values, and hence also amplifies the noise. Consequently, (i) the noise power is greater than for the case without an equalizer at the detector input, and (ii) the noise is closed, as shown in fig. 2.21.

The relation between $\hat{\Xi}$, the Fourier transform of $\eta(t)$ and the Fourier transform $\Xi(e^{j\omega T_S})$ of the sample ACF ζ_i is as follows –

$$\Xi(e^{j\omega T_S}) = \frac{1}{T_S}\sum_{n=-\infty}^{\infty}\left|\hat{\Xi}\left(\omega + \frac{2\pi n}{T_S}\right)\right|^2, |\omega| \le \frac{\pi}{T_S} \quad ...(v)$$

At the detector, the noise power is as follows –

$$\sigma^2_{n-LE-ZF} = N_0 \frac{T_S}{2\pi}\int_{-\pi/T_S}^{\pi/T_S} \frac{1}{\Xi(e^{j\omega T_S})} d\omega \quad ...(vi)$$

Only when the spectral density Ξ has no singularities, it is finite.

(ii) The Mean-square Error Criterion – Ultimately, the aim of an equalizer is minimization of the bit error probability, not of the inter symbol interference reduction. The zero-forcing equalizer is ill-suited for this purpose due to noise enhancement. The minimization of the mean square error (MSE) reduction between the output of the equalizer and the transmit signal is a good criterion.

Hence, we search for a filter that minimizes –

$$\text{MSE} = E\{|\varepsilon_i|^2\} = E\{\varepsilon_i \varepsilon_i^*\} \quad ...(vii)$$

This is obtained with a filter whose coefficients $\mathbf{e}_{opt}$ are expressed as follows –

$$\mathbf{e}_{opt} = \mathbf{R}^{-1}\mathbf{p} \quad ...(viii)$$

where $\mathbf{p} = E\{\mathbf{u}_i^* c_i\}$ denotes the cross-correlation between the transmit signal and the received signal, $\mathbf{R} = E\{\mathbf{u}_i^* \mathbf{u}_i^T\}$ denotes the correlation matrix of the received signals. The concatenation of the noise-whitening filter with the equalizer E(z), considering the frequency domain, contains the transfer function as follows –

$$\tilde{E}(z) = \frac{1}{\Xi(z) + \frac{N_0}{\sigma_s^2}} \quad ....(ix)$$

which is the Wiener filter's transfer function. Then, the mean square error is as follows –

$$\sigma^2_{n\text{-LE-MSE}} = N_0 \frac{T_S}{2\pi}\int_{-\pi/T_S}^{\pi/T_S} \frac{1}{\Xi(e^{j\omega T_S}) + \frac{N_0}{\sigma_s^2}} d\omega \quad ...(x)$$

The noise power of an minimum mean square error equalizer is lesser as compared to a zero-forcing equalizer by comparing with equation (vi), as shown in fig. 2.21.

Q.26. *Discuss multilevel coding.* ***(R.G.P.V., Dec. 2016)***

Ans. The idea of multilevel coding or generalized code concatenation is that m outer codes $C^{(i)}$, i = 1, ... m protect the labeling of a partitioned signal constellation B. Fig. 2.22 shows the structure of such a multilevel encoder. The block DEMUX demultiplexes the binary vector $\vec{b} = (b_1, ..., b_m)^T$ into its components b_i which are assigned to separate levels of the MLC. In general multilevel coding techniques, the level codes $C^{(i)}$ can have different complexity and rate.

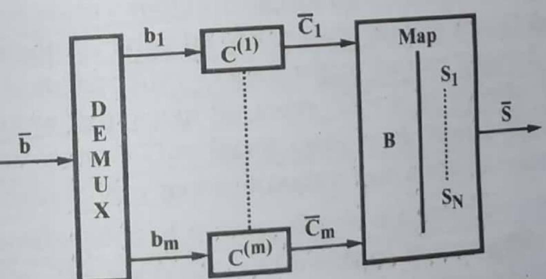


Fig. 2.22 Structure of a Multilevel Encoder

Here we restrict ourselves to a special case of multilevel coding, using only rate 1/N binary repetition codes in each level. Thus, each bit b_i is repeated N times to a codeword $\vec{c_i} = (c_{i,1}, ... c_{i,N})^T$ and the codewords $\vec{c_i}$, i = 1, ... m build a m × N codeword matrix $\overline{\overline{C}}$. Each column i of the codeword matrix $\overline{\overline{C}}$

is mapped to one point s_i from a complex valued signal constellation S_i with cordinality $|S_i| = 2^m$. Therefore, each binary codeword matrix $\overline{\overline{C}}$ gives a codeword $\overline{S} = (S_1 \ldots S_N)^T$ of length N in the N-dimensional complex space, $s_i \varepsilon S_i$, i = 1 ..... N. This result is so-called polyalphabetic codes and we obtain the code word matrix $\overline{\overline{C}}$ –

$$\overline{C} = \begin{pmatrix} \overline{C}_1^T \\ \vdots \\ \overline{C}_m^T \end{pmatrix} = \begin{pmatrix} C_{1,1} & \cdots & C_{1,N} \\ \vdots & \ddots & \vdots \\ \underbrace{C_{m,1}}_{s_1 \varepsilon S_1} & \cdots & \underbrace{C_{m,N}}_{s_N \varepsilon S_N} \end{pmatrix}$$

Instead of changing the signal constellation for every column i, it is also possible to use a fixed constellation B for all columns of the matrix $\overline{\overline{C}}$.

OO

UNIT 3

MULTI ANTENNA SYSTEM

SMART ANTENNAS

Q.1. *Explain in detail the principle working of smart antenna.* **(R.G.P.V., June 2014)**

Or

Write short note on smart antenna. **(R.G.P.V., Dec. 2011)**

Or

Explain the operating principle of smart antennas. **(R.G.P.V., June 2016)**

Or

Discuss the principle working of smart antennas. **(R.G.P.V., June 2017)**

Ans. Smart antennas are antennas with multiple elements, where signals from distinct elements are created or combined through an adaptive algorithm, when the smart antenna is at the transmitter end and at the receiver end respectively.

Imagine you are in a room sitting in a chair. Someone is talking to you in the room and while speaking he or she starts moving around the room. Your ears and brain have the ability to track where the user's speech is originating from as they move throughout the room. This is very similar to how smart antenna systems operate. They locate users, track them, and exhibit optimal RF signals to them as they move through a base station's coverage area.

Usually, combination of antenna signals is a linear combination using a weight vector **w**. The approaches of obtaining **w** necessarily distinguishes smart antenna systems. There exists a strong relationship between diversity systems and multiantenna systems. In fact, an RX having antenna diversity is a smart antenna.

Q.2. *Give the advantages of smart antenna.*

Or

Describe smart antenna advantages. **(R.G.P.V., May 2018)**

Ans. The advantages of smart antenna are as follows –

(i) Increased Number of Users – Due to the targeted nature of smart antennas frequencies are reused permitting an increased number of users.

(ii) Increased Bandwidth – The bandwidth available increases form the reuse of frequencies and also in adaptive arrays as they can utilize the many paths which a signal may follow to reach a device.

(iii) Reduced Interference – Interference which is usually caused by transmissions which radiate in all directions is less likely to take place because the directionality introduced by the smart antenna. This aids both the ability to reuse frequencies and achieve larger range.

Q.3. Write the drawbacks of smart antenna. ***(R.G.P.V., Dec. 2016)***

Ans. Drawbacks of smart antenna are as follows –

(i) Smart antennas are far more complicated compare to the traditional antennas.

(ii) Because smart antennas are extremely complex, utilizing the latest in processing technology, they are far more expensive compare to the traditional antennas.

(iii) Due to the antenna arrays which are utilized by smart antenna systems, they are much larger in size compare to the traditional antenna systems.

Q.4. Why smart antenna are required ?

Ans. Smart antennas are used for following purposes –

(i) Capacity Increase – The SIR (signal-to-interference ratio) can be improved by using smart antenna, which permits the number of users in the system to be maximized.

(ii) Link Quality Improvement – The transmission quality on each single link can be improved, by increasing signal power and/or reducing interference power.

(iii) Coverage Increase – It is assumed that the smart antenna is at the receiver (RX). Now, the receiver (RX) can form an antenna pattern in the transmitter (TX) direction, when the spatial position of the transmitter is known. This results in higher receive power.

Q.5. Give the basic differences of smart antennas over the normal antennas. ***(R.G.P.V., June 2016)***

Ans. The basic differences of smart antennas over normal antennas are given below –

(i) Smart antennas are more directive than traditional sectorized or omnidirectional antennas.

(ii) A smart antenna transceiver is much more complicated than a traditional base station transceiver.

(iii) Using smart antennas, an increase of the range of coverage by a base station is possible since they are able to focus their energy toward the intended users instead of directing and wasting it in other unnecessary directions.

(iv) Smart antenna base stations are much more expensive than conventional base stations.

Also refer to the ans. of Q.4.

Q.6. What do you mean by antenna array ? Does antenna array system require for the communication purposes ? ***(R.G.P.V., June 2016)***

Ans. A common method of combining the radiations from a group or array of similar antennas in which the phenomena of wave interference is involved is known as antenna array. The total field produced by an antenna array system at a great distance from it, is the vector sum of the fields produced by the individual antennas of the array system. If the individual antennas of the array are equally spaced along a straight line then antenna array is said to be linear. Individual antennas of an antenna array system is also termed as elements. Hence, linear antenna array is a system of equally spaced elements. Since antennas may be put in various configuration like in straight lines, circles, triangles, rectangles and hence there are large number of possible configuration. However experience shows that in practices only limited number of antenna arrays are of practical use.

Q.7. How can we increase the capacity by using smart antennas ?

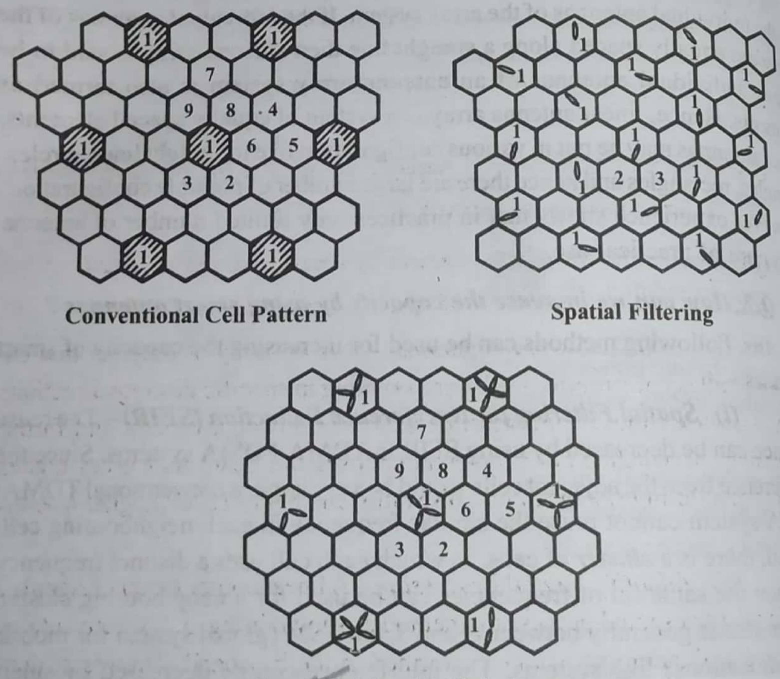
Ans. Following methods can be used for increasing the capacity of smart antennas –

(i) Spatial Filtering for Interference Reduction (SFIR) – The reuse distance can be decreased by using SFIR in TDMA/FDMA systems. Since the interference from the adjacent cells would be very large, a conventional TDMA/FDMA system cannot reuse the similar frequency in each neighbouring cell. Instead, there is a ***cluster*** of cells, in which each cell uses a distinct frequency and then the same set of frequencies can be used for a neighbouring cluster. Cluster size is generally between 3 and 7 for GSM (global system for mobile communications) like systems. The interference can be decreased by smart antennas. Therefore, the cells with the similar frequency can be put closer together. The cluster size is thus reduced, which causes an immediate improvement in total spectral efficiency. The number of users per area can be increased proportionately, by using same frequencies in more cells.

(ii) Space Division Multiple Access (SDMA) – In this technique the cluster size is constant, while the number of users in a given cell is maximised. The base station distinguishes various users by their distinct spatial signatures, therefore multiple users can be served on the same frequency/time slot.

(iii) Capacity Increase in CDMA Systems – The capacity of CDMA systems can also be improved using smart antennas. This capacity increase obtained is subtly distinct from TDMA/FDMA systems. The same carrier frequency is used by all mobile stations (MSs) in a CDMA system and they are distinguished by the spreading codes used by them. The suppression of users is imperfect with other codes. When the residual interference from all the other K users is comparable with the admissible SIR, then a cell is decided to be full. The SIR is as follows –

$$SIR_{threshold} = \frac{P_{desired}}{\Sigma P_{interfere}} \quad ...(i)$$



Conventional Cell Pattern **Spatial Filtering**

Space Division Multiple Access

Fig. 3.1

where, $\Sigma P_{interfere} = K.P_{interfere}$ is the case of perfect power control. The number of users is of the order of 30 per cell for voice CDMA systems. Therefore, a

CDMA system has a large number of possible interferers, each provides only a small interference part. Since the number of interferers that can be suppressed is smaller than the number of antenna elements, therefore it is impossible to suppress interference. Instead, the object of smart antenna is to improve the received signal power. The received signal power can be increased by an N_r-element antenna by a factor of N_r, therefore –

$$P_{desired} = M_c . N_r . P_{interfere} \quad ...(ii)$$

where M_c denotes the spreading gain. In that condition, the number of admissible users in the cell is given as –

$$K = \frac{M_c . N_r}{SIR_{threshold}} \quad ...(iii)$$

and hence linearly increases with the number of antenna elements.

(iv) Capacity Increase in Third-generation CDMA Systems – High-rate data transmission is considered an important application in third-generation CDMA systems. A high-rate data user creates more interference because of the lower spreading factor, therefore it needs to be supressed by placing a null in its direction.

Q.8. What is spatial division multiple access ? **(R.G.P.V., Dec. 2016)**

Ans. Refer to the ans. of Q.7 (ii).

Q.9. Discuss the receiver structures of smart antenna in detail.

Ans. The receiver structures of smart antennas are as follows –

(i) Switched-beam Antennas – It is an array of antenna that can make a small set of patterns, i.e. beams pointing in specific discrete directions. Then, a switch chooses one possible beam for downconversion and further processing. A switched-beam antenna chooses the beam that gives highest SINR (signal-to-interference-and-noise ratio) and highest SNR (signal-to-noise ratio).

(ii) Adaptive Spatial Processing – A linear combination of signals is used in adaptivespatial processing method as shown in fig. 3.2, where antenna weighting and summing can be performed in the baseband. Therefore, there are no limitations with respect to weights. They can be adapted according to the curent channel state. The disadvantage of this method is that it requires N_r complete downconversion chains for N_r antenna elements, and hence considerably more expensive, and also consumes more power.

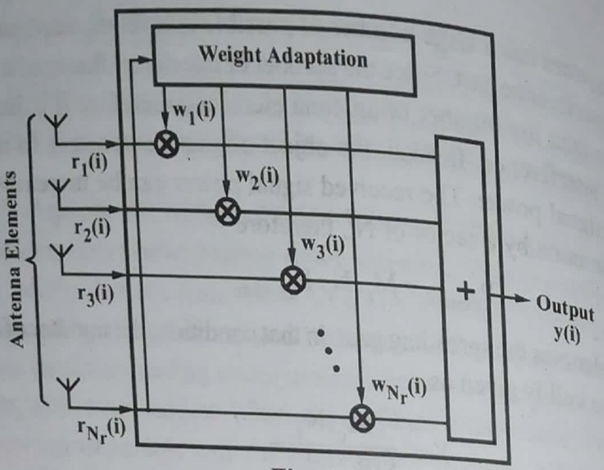


Fig. 3.2

(iii) Adaptive Space-time Processing – Adaptive space-time processing method has to be used to fully exploit the possibilities of distinct multipath components (MPCs). A two-dimensional Rake receiver is the optimum linear RX, which weights all resolvable MPCs and adds them up. Such a space-time processor is shown in fig. 3.3. The total processor has N_rL weights when the Rake receiver has L taps. The output from that processor is transmitted on to a decoder.

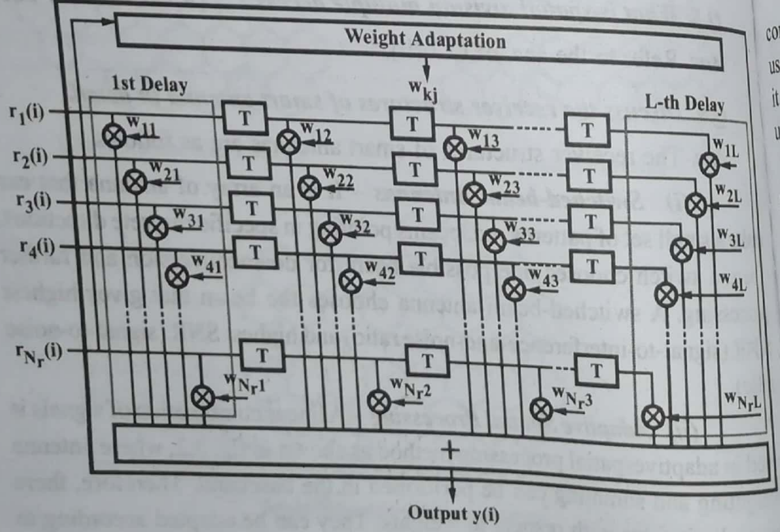


Fig. 3.3 Space-time Filter Principle

(iv) Space-time Detection – Space-time processing and decoding/detection have to be performed jointly for optimum reception. For detection, the special properties of the received signal, like the finite-alphabet properties need to be considered. The optimum detector is a generalized MLSE (maximum likelihood sequence estimation) receiver. These structures are quite complicated, and the performance gain compared to linear processing, does not justify the extra effort. That is why, they are not used in often practice.

Q.10. Write down the algorithms for adaptation of antenna weights.

Ans. These algorithms can be classified into spatial reference and temporal reference, which are discussed below –

(i) Spatial Reference Algorithms – The antenna attempts to make a beam pattern that introduces maxima in the direction of main DOAs (direction of arrivals) of the required user, and nulls in the direction of MPCs coming from interferers in the SR (spatial reference) algorithms. An spatial reference algorithm works in the following manner –

(a) Firstly, it computes the DOAs, ϕ_n, of the multipath components. The DOAs show small variations, over frequency and time, which is the main benefit of using them. The signals at the antenna elements are directly observable. Hence, it is essential to remove the DOAs from these signals. This can be performed by the high-resolution algorithms and spatial Fourier transformations.

(b) The next step is to associate DOAs with specific users. During communications, the MPCs can originate from distinct users – the required user and also interferers. Hence, it is essential to separate them. For this purpose, it is essential to have a training sequence, or some other characteristic that is unique for each user.

(c) It is possible to make the beam pattern, that increases the SINR, after the DOAs are established for user and interferer. This beam pattern is used for actual data reception.

(ii) Temporal Reference Algorithms – Temporal reference (TR) algorithms use a training signal, which acts as a ***temporal reference***, and is matched against the output from the smart antenna. Antenna weights **w** are selected in such a way that deviation between training sequence and combiner output is reduced. We can use the SIR, the BER (bit error rate) during the training sequence, the MMSE (minimum mean square error), or other appropriate criterion as a criterion for deviation. A TR (temporal reference) algorithms works in the following manners (fig. 3.4) –

(a) The signal at all antenna elements can be received by the smart antenna during training phase, and deviation from the known signal is reduced by adjusting the weights in an appropriate way.

(b) The antenna weights remain constant during transmission of user data. They are used to weight incoming signals before they are combined and decoded/demodulated.

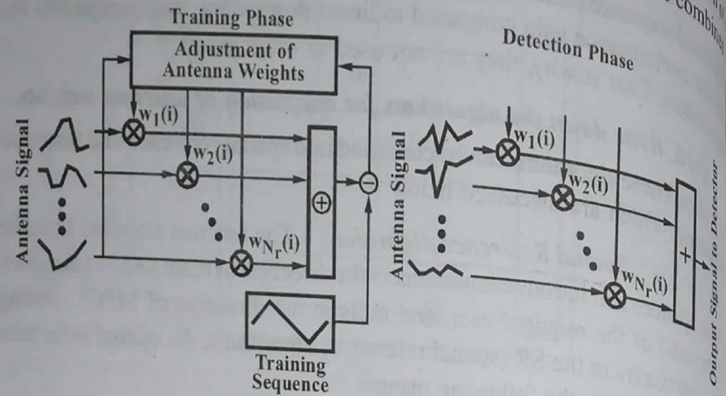


Fig. 3.4 Temporal Reference Algorithms Principle

Q.11. Give an overview of direction of arrival algorithms.

(R.G.P.V., Dec. 2016)

Ans. Refer to the ans. of Q.10.

Q.12. Explain the behaviour of smart antenna during uplink and downlink.

Ans. Behaviour of smart antennas during uplink and downlink is discussed below for following two cases –

(i) Time Domain Duplexing – In time domain duplexing, the channel impulse response is similar for uplink and downlink when there is no movement of either the TX or the RX i.e., in a static environment. The transfer function from the mth base station (BS) antenna to the mobile station (MS) antenna is similar to the transfer function from the mobile station antenna to the mth base station antenna, $(h_d)_m$. Hence, when we select suitable antenna weights in such a way that they constructively add signals from the antenna elements during the uplink, the same antenna weights will also make sure that during the downlink signals from distinct base station antenna elements will add up constructively at the mobile station antenna. Also, the base station will cause small or zero interference to the other mobile stations (MSs) in the downlink transmission phase, when base station antenna weights are selected to supress interference from those mobile stations (MSs) during the uplink.

However, the interference for the uplink and downlink is not perfectly reciprocal. The interference seen by the MS originates only from other base stations, not from other mobile stations (MSs). Although, the antenna weights W which are computed during the uplink do not consider the above situation, because the required base station cannot see these other BSs. Hence, interference supression in the downlink is less effective as compared to that in the uplink.

(ii) Frequency Domain Duplexing – The uplink and downlink occur at distinct frequencies in frequency domain duplexing system. The frequency separation is of the order of 100 MHz for cellular system. This is much higher as compared to the channel coherence bandwidth. As a result, small-scale fading of the channel impulse response in the uplink and downlink are decorrelated totally. Alternately, superposition of MPCs with distinct phases produces small-scale fading. The relative phaseshifts are dependent on the carrier frequency among other things. When duplexing distance is large, phaseshifts are adequately distinct in uplink and downlink that superposition of MPCs happens in distinct manner.

Therefore, the small-scale fading and thereby the instantaneous impulse response of the channel differ for uplink and downlink.

Q.13. Discuss in detail the multiuser diversity and random beam forming of smart antennas.

Ans. The multiuser diversity and random beam forming of smart antennas are discussed below –

(i) Scheduling and Multiuser Diversity – Consider the downlink in a system where the base station and mobile stations have only a single antenna, and multiple mobile stations communicate with one base station. All links experience flat Rayleigh fading. It is assumed that the mean power is identical for all users for the moment. Also consider a conventional TDMA system, therefore each user is allocated a fixed timeslot, and communicates during that timeslot. Then, the probability for user k to experience a specific SNR γ is given by the exponential distribution given below –

$$Pr(r_k \leq \gamma) = 1 - e^{\left(-\frac{\gamma}{\bar{\gamma}}\right)} \qquad ...(i)$$

This probability neither depend on the user k nor on the number of users K. The round-robin is the best method we can apply when the transmitter (BS) does not know anything regarding the channel. It is clear that, it treats users in such a way that each user get the same amount of time to communicate.

However, better results can be obtained when the base station knows the channel state information for the downlink of all users. At any given time, chances are that there are bad links to some users, while there is a very good link to other users. Now, the optimum strategy for the base station is to communicate always with the user who has the best link, that is the highest intantaneous path gain, and hence the highest SNR γ. The distinct users k become ***instantaneously best*** when the channel changes. The cdf(cumulative distribution function) for the active users is similar as for selection diversity. Hence, multiple users work like the diversity branches. Although, there is no need of additional antennas to attain that diversity. Hence, this technique is known as multiuser diversity. It is optimum because the capacity is maximized.

In a typical cellular system, users that are far away from the base station have a worse SNR compared to users that are very close. Hence, the problem with the above technique is that the MS nearest to the BS can use the channel most often, on the other hand MSs that are far away are almost never ***instantaneous best***, and hence are selected much less frequently. As a result, the data rate for users near the BS is much greater than for users at the cell boundary. This problem can be overcome by proportional scheduling. The base station does not communicate with the absolutely best user, but with the user whose ratio of instantaneous to average SNR $\gamma/\bar{\gamma}$ is the highest. Therefore, each user is allocated the similar amount of time when the normalized fading statistics of the distinct users are same i.e., all the users are Rayleigh-fading.

The main requirement for scheduling is that the base station has knowledge of the instantaneous downlink channel.

(ii) Delay Considerations and Random Beamforming – The inherent delay is the another important problem of multiuser diversity. The base station retains communication with the selected user for suitably one coherence time of the channel, when it always uses proportionally fair scheduling. The coherence time is large, when user mobility is low. It should be noted that each user has to wait until its $\gamma/\bar{\gamma}$ becomes instantaneously highest and it can communicate. This waiting time is approximately KT_{coh}. A system having large coherence times and many users shows large waiting time, which is prohibited.

This problem can be solved by using multiple antenna at the base station, and this is known as ***random beamforming***. The time varying complex coefficients are multiplied with the signals transmitted from the distinct antennas of the base station. This can be interpreted in two manners as follows –

(a) The combination of physical channel and multiple antennas can be considered as an equivalent channel. Coherence time T_{coh} can be reduced by varying the coefficients for antenna elements, which in turn enforces time variations on the channel.

(b) Each vector of transmit weight coefficients corresponds to a beam pattern. Irrespective of the type, a beam always improves the channel to some MS. Then scheduling guarantees that the improved mobile station (MS) is selected for communication. When the coefficients are changed, the beam pattern changes, and a distinct mobile station gets improved.

MULTIPLE INPUT MULTIPLE OUTPUT SYSTEMS, MULTIUSER MIMO

Q.14. What do you mean by MIMO systems ? How does spatial multiplexing work ?

Ans. The multiple-input multiple-output (MIMO) systems are also known as multiple element antenna (MEA) systems. These are wireless systems with MEAs at both link ends.

The MEAs of a MIMO is used for following three distinct purposes –

(i) Diversity

(ii) Beamforming

(iii) Spatial multiplexing.

The first two methods are similar as for smart antennas while, spatial multiplexing permits direct enhancement of capacity through simultaneous transmission of multiple datastreams.

The MEAs are used by spatial multiplexing at the transmitter for transmission of parallel datastreams as is shown in fig. 3.5. An original high-rate datastream is multiplexed into many parallel datastreams, and each of these is sent from one transmit antenna element. These datastreams are mixed-up by the channel, therefore each of the receive antenna elements obtains a combination of them. The received signals shows ***linearly independent*** combinations, when the channel is working properly. In this case, the datastreams can be separated by the suitable signal processing at the receiver. A fundamental condition is that the number of receive antenna element is greater than or equal to the number of transmit datastreams. Obviously, this technique permits the data rate to be increased by a large factor of min (N_t, N_r).

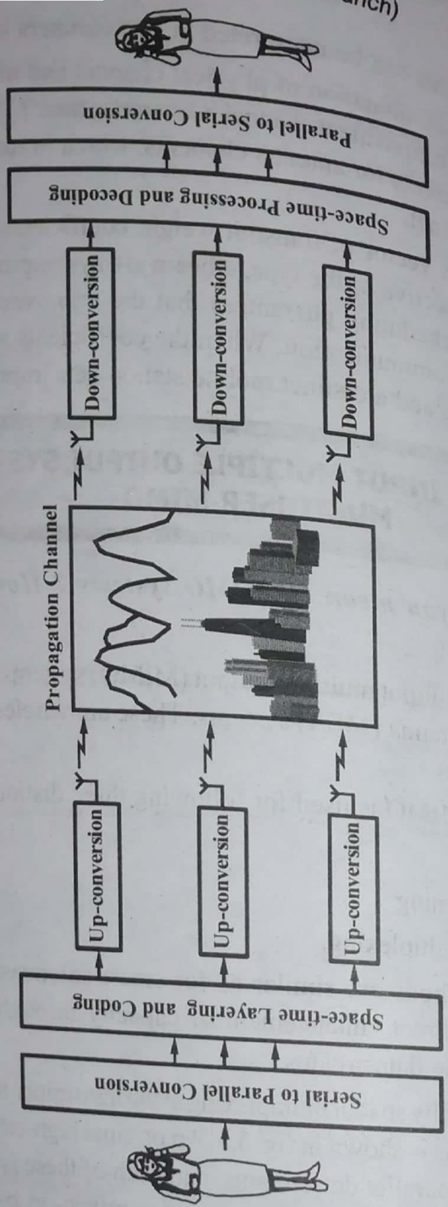


Fig. 3.5 Spatial Multiplexing Principle

Q.15. Write short notes on MIMO. ***(R.G.P.V., June 2016)***

Ans. Refer to the ans. of Q.14.

Q.16. Discuss MIMO system model. ***(R.G.P.V., Dec. 2016)***

Ans. A system model block diagram of multiple-input multiple-output is shown in fig. 3.6. The datastream enters an encoder at the transmitter. The outputs of encoder are send to N_t transmit antennas. The signal is then

transmitted by the wireless propagation channel from the antennas, which is supposed to be quasi-static and frequency-flat when not stated otherwise. Here the meaning of quasi-static is that the channel coherence time is too large that a large number of bits can be send with this time.

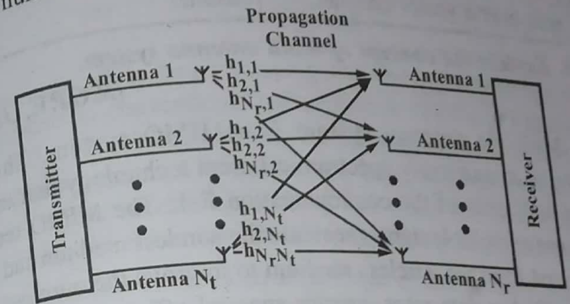


Fig. 3.6

The $N_r \times N_t$ matrix of the channel can be represented as –

$$\mathbf{H} = \begin{pmatrix} h_{11} & h_{12} & \dots\dots & h_{1N_t} \\ h_{21} & h_{22} & \dots\dots & h_{2N_t} \\ \vdots & \vdots & \dots\dots & \vdots \\ h_{N_r1} & h_{N_r2} & \dots\dots & h_{N_rN_t} \end{pmatrix} \quad \text{...(i)}$$

In this matrix, h_{ij}'s denotes the attenuations from the jth transmit to the ith receive antenna.

The received signal vector given by,

$$\mathbf{r} = \mathbf{Hs} + \mathbf{n} = \mathbf{x} + \mathbf{n} \quad \text{...(ii)}$$

is received through N_r antenna elements, where n dentoes the noise vector and s denotes the transmit signal vector.

Q.17. Describe MIMO in detail. ***(R.G.P.V., May 2018)***

Ans. Refer to the ans. of Q.14 and Q.16.

Q.18. Write down the comparison between smart antenna and MIMO technology.

Ans. Comparison between smart antenna and MIMO technology are given below –

S.No.	*Smart Antenna*	*MIMO Technology*
(i)	In smart antenna, only TX or RX or both are equipped with more than one antenna.	In MIMO, both sides of the communication have more than one antenna.

(ii)	Some smart antenna performs better in line-of-sight (LOS) or close to line-of-sight system.	Multi input multi output technology can perform well in non line-of-sight.
(iii)	In smart antenna, the data is sent over a vector channel.	In MIMO, data is sent over matrix channel.

Q.19. Explain the concept of multi antenna system.

(R.G.P.V., June 2016)

Ans. Multiple antenna system is a MIMO system. The MIMO transmission is an extremely spectrum-efficient technology that uses many antennas at both ends of the communication link. The MIMO technology exploits the use of multiple signals sent into the wireless medium and multiple signals received from the wireless medium to improve the wireless channel performance. Higher data rates, greater spectral efficiency, greater range, an increased number of users and enhanced reliability can be obtained by the MIMO technology.

Q.20. How does a smart antenna differ from multiple input multiple output system. Give their principle of working in detail.

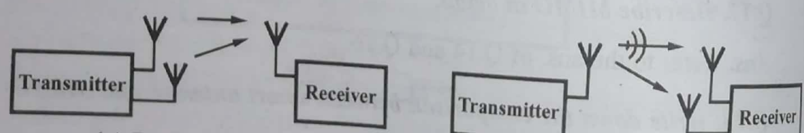
(R.G.P.V., June 2013)

Or

Discuss the principle and working of smart antennas used in communication systems. How they are different from MIMO systems ?

(R.G.P.V., June 2015)

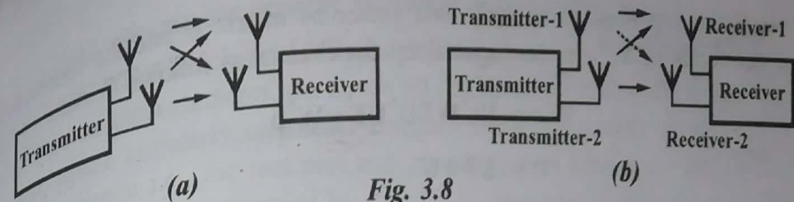
Ans. The smart antennas enhance conventional, one-dimensional radio systems. The most common smart-antenna systems employ beam forming or transmit diversity to concentrate the signal energy on the main path is shown in fig. 3.7 (a) and receive the combination to capture the strongest signal at any given moment is shown in fig. 3.7 (b).



(a) Beam Forming ***(b) Diversity***

Fig. 3.7

It should be noted that beam forming and receive combining are only multipath mitigation methods and do not multiply data throughout over the wireless channel. Fig. 3.8 (a) shows the both combined together demonstrate an ability to improve performance incrementally in point-to-point applications.



(a) ***Fig. 3.8*** ***(b)***

Although, while beam forming and receive combining are valuable enhancements to conventional radio systems, MIMO is a paradigm shift as shown in fig. 3.8 (b), dramatically changing perceptions of and responses to multipath propagation. The transmitted information by both the antennas is different. While receive combining and beam forming increase spectral efficiency by one or two b/s/Hz at a time, MIMO multiplies the b/s/Hz.

The powerful effect of smart antenna is that in the presence of random fading caused by multipath propagation, the probability of losing the signal vanishes exponentially with the number of decorrelated antenna elements being employed. Obviously, in an MIMO link, the advantages of conventional smart antennas are retained since the optimization of the multi-antenna signals is carried out in a larger space, therefore giving additional degrees of freedom. In particular, MIMO systems can give a joint transmit-receive diversity gain, as well as an array gain upon coherent combining of the antenna elements.

Also refer to the ans. of Q.1.

Q.21. Derive the expression of the capacity for MIMO system in non-fading channels.

Ans. Let us begin with the capacity equation (Foschini's equation) for normal AWGN(additive white Gaussian noise) channels. The information-theoretic capacity of such a channel as given by Shannon is as follows –

$$C_{shannon} = \log_2(1 + \gamma.|\mathbf{H}|^2) \quad ...(i)$$

where, H denotes the normalized transfer function from the transmitter to the receiver and γ denotes the SNR at the receiver. Equation (i) shows that the capacity of MIMO system increases logarithmically with the SNR, therefore, boosting the transmit power is not a much effective way of increasing capacity.

Now, consider the MIMO case, where matrix represents the channel. Then consider a singular value decomposition of the channel is as follows –

$$\mathbf{H} = \mathbf{W\Sigma U^{\dagger}} \quad ...(ii)$$

where $\mathbf{W}$ and $\mathbf{U}^{\dagger}$ denotes the unitary matrices composed of the left and right singular vectors respectively, and Σ denotes the diagonal matrix containing singular values. Then, the received signal is as follows –

$$\mathbf{r} = \mathbf{Hs} + \mathbf{n} \quad ....(iii)$$

$$= \mathbf{W\Sigma U^{\dagger}s} + \mathbf{n} \quad ....(iv)$$

By multiplying the transmit data vector by matrix **U** and the received signal vector by $\mathbf{W}^\dagger$ we can diagonalizes the channel as follows –

$$\mathbf{W}^\dagger\mathbf{r} = \mathbf{W}^\dagger\mathbf{W}\Sigma\mathbf{U}^\dagger\mathbf{U}\tilde{\mathbf{s}} + \mathbf{W}^\dagger\mathbf{n}$$

$$\tilde{\mathbf{r}} = \Sigma\,\tilde{\mathbf{s}} + \tilde{\mathbf{n}} \quad ...(v)$$

where, Σ denotes the diagonal matrix with R_H nonzero entries σ_k, R_H denotes the rank of **H**, and σ_k denotes the kth singular value of **H**. Consequently, we have R_H parallel channels, and obviously the parallel channels capacity just adds up.

Hence, the capacity of channel **H** is given by the sum of the subchannels capacities as follows –

$$C = \sum_{k=1}^{R_H} \log_2\left[1 + \frac{P_k}{\sigma_n^2}\sigma_k^2\right] \quad ...(vi)$$

where, P_k = Power assigned to the kth eigenmode

σ_n^2 = Noise variance.

It is assumed that $\Sigma P_k = P$ is not dependent on the number of antennas. Then capacity expression will be as follows –

$$C = \log_2\left[\det\left(\mathbf{I}_{N_r} + \frac{\bar{\gamma}}{N_t}\mathbf{H}\mathbf{R}_{SS}\mathbf{H}^\dagger\right)\right] \quad ...(vii)$$

where $\bar{\gamma}$ denotes the mean SNR per RX branch, $\mathbf{R}_{SS}$ denotes the correlation matrix of the transmit data and $\mathbf{I}_{Nr}$ denotes the $N_r \times N_r$ identity matrix.

Full-channel-state Information at the Receiver and No-channel-state Information at the Transmitter – When no-channel-state information is available at the transmitter, but the receiver knows the channel accurately, then it is optimum to allocate equal transmit power to all TX antennas $P_k = P/N_t$, and employ uncorrelated datastreams. Hence, the capacity can be given as follows –

$$C = \log_2\left[\det\left(\mathbf{I}_{N_r} + \frac{\bar{\gamma}}{N_t}\mathbf{H}\mathbf{H}^\dagger\right)\right] \quad ...(viii)$$

Irrespective of whether the channel is known at the transmitter or not, the capacity of a MIMO system increases linearly with min (N_t, N_r).

Full-channel-state Information at the Receiver and Full-channel-state Information at the Transmitter – In this case, it will be more advantageous to allocate power between the distinct transmit antennas depending upon the channel state information.

Q.22. Discuss the capacity of MIMO in flat-fading channels.

Ans. When the channel is Rayleigh-fading, and fading is not dependent at distinct antenna elements, the h_{ij} are iid (independent identically distributed) zero mean, circularly symmetric complex Gaussian random variables with unit variance i.e., the real part and imaginary part each has variance 1/2. Therefore, the power carried by each h_{ij} is chi-square-distributed with two degrees of freedom. It needs the presence of ***heavy multipath*** i.e., several multipath components (MPCs) of approximately equal strength and also a enough distance between the antenna elements. There is a high probability that the channel matrix is full rank and the eigen values are identical to each other, because the fading is independent. Therefore, with the increase of number of antenna elements capacity increases linearly. In this way, the presence of heavy multipath becomes an important benefit in MIMO systems.

There are two distinct definitions for capacity of MIMO systems, which are as follows –

(i) Shannon (Ergodic) Capacity – This is the predicted value of the capacity, considering over all channel realizations. Shannon capacity considers an infinitely long code which extends over all the distinct channel realizations.

(ii) Outage Capacity – This is the minimum transmission rate which is achieved over a specific fraction of time i.e., 90% or 95%. The data are encoded with a near-Shannon-limit-achieving code which extends over a period that is much shorter compared to the channel coherence time.

Perfect Channel-state Information at the Receiver and No Channel-state Information at the Transmitter – Fig. 3.9 shows the capacity variation for few important systems in a fading channel without CSI. The exact expression for the Shannon capacity is as follows –

$$E\{C\} = \int_0^\infty \log_2\left[1 + \frac{\bar{\gamma}}{N_t}\lambda\right]\sum_{k=0}^{m-1}\frac{k!}{(k+n-m)!}$$

$$[L_k^{n-m}(\lambda)]^2\lambda^{n-m}e^{-\lambda}d\lambda \quad ...(i)$$

where n denotes the max (N_t, N_r), m denotes the min (N_t, N_r) and $L_k^{n-m}(\lambda)$ denotes the Laguerre polynomials.

When $N_t \geq N_r$ the upper and lower limits for capacity distribution are as follows –

$$\sum_{k=N_t-N_r+1}^{N_t} \log_2\left[1 + \frac{\bar{\gamma}}{N_t}\chi_{2k}^2\right] < C < \sum_{k=1}^{N_t}\log_2\left[1 + \frac{\bar{\gamma}}{N_t}\chi_{2N_r}^2\right] \quad ...(ii)$$

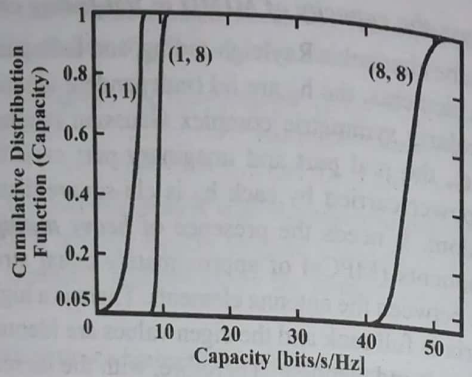


Fig. 3.9

where χ^2_{2k} denotes the chi-square-distributed random variable with 2k degrees of freedom. The lower limit corresponds to capacity, which can be obtained by a BLAST system. The upper limit corresponds to the condition when there is a separate array of receive antennas for each transmit antenna. It receives the signal in such a way that there is no interference from other transmit streams.

Perfect Channel-state Information at the Transmitter and Receiver – When the transmitter has full channel knowledge, it can perform beamforming and direct the energy in a better way toward the receive array. Hence, the capacity and SNR will be improved by increasing the number of transmitter antennas.

Q.23. Explain the impact of the channel correlation on MIMO systems.

Ans. The capacity of a MIMO system is decreased by correlation of the signals at distinct antenna elements. Correlation is affected by the arrangement and spacing of antenna elements as well as the angular spectrum of the channel. A uniform angular power spectrum leads to a decorrelation of incident signals for antennas that are spaced half a wavelength apart. An increase in correlation is provided by a smaller angular spread of the channel. The Kronecker model assumes that correlation at the transmitter does not depend on correlation at the receiver. Then, the channel matrix realization is obtained as follows –

$$\mathbf{H}_{kron} = \frac{1}{\sqrt{tr(\mathbf{R}_{RX})}} \mathbf{R}_{RX}^{1/2} \mathbf{G}_G \mathbf{R}_{TX}^{1/2}$$

where, $\mathbf{G}_G$ denotes the matrix with iid complex Gaussian entries with unit variance.

Q.24. Explain the line-of-sight effect versus non-line-of-sight for MIMO systems.

Ans. In some cases, there are different fading statistics resulting due to LOS (line of sight) connection between transmitter and receiver. The fading statistics of an SISO link is Ricean instead of Rayleigh. The channel matrix for a MIMO system can be defined as follows –

$$\mathbf{H} = \sqrt{\frac{K_{LOS}}{K_{LOS}+1}}\hat{\mathbf{H}}_{LOS} + \sqrt{\frac{1}{K_{LOS}+1}}\hat{\mathbf{H}}_{res} \quad ...(i)$$

where $\hat{\mathbf{H}}_{res}$ has zero-mean Gaussian entries, $\hat{\mathbf{H}}_{LOS}$ denotes the purely deterministic matrix, and K_{LOS} denotes the ratio of powers in the LOS to those in residual components. When the distance between the transmitter and receiver is greater than the Rayleigh distance, then LOS generates a matrix $\hat{\mathbf{H}}_{LOS}$ of rank 1. It means that the singular value spread of matrix equation (i) is greater than for a NLOS (non line of sight) matrix. Therefore, if assuming equal SNR, the capacity of a LOS channel is lower compared to the NLOS channel. Although, it should be noted that, the SNR is often better for LOS case as compared to the NLOS case.

When the LOS component is a plane wave, then a strong line-of-sight component causes to a greater spread of eigenvalues. When antenna elements are suitably spaced, then a spherical wave causes a transfer function matrix of full rank.

Q.25. Discuss the Horizontal BLAST (H-BLAST) with block diagram.

Ans. The simplest possible layered space-time structure is Horizontal Bell labs LAyered Space Time (H-BLAST), which is shown in fig. 3.10. Firstly, the datastream is demultiplexed into N_t parallel streams by the transmitter and each of which is encoded separately. Then, the transmission of each encoded datastrem is done from a different transmit antenna. Various datastreams are mixed-up by the propagation channel. The RX separates datastreams out by interference and nulling subtraction. In other words, the receiver proceeds in the following steps –

(i) It considers the first datastream as the useful one. The other datastreams are considered as interference, which can be supressed by optimum combining. The antenna elements available at the RX are $N_r \geq N_t$. It can receive the first datastream with better quality, when $N_r > N_t$. When $N_r = N_t$, it can suppress all $N_t - 1$ interfering datastreams, and the required datastream is received with diversity order 1. However, in both cases, interference from the other datastreams are removed.

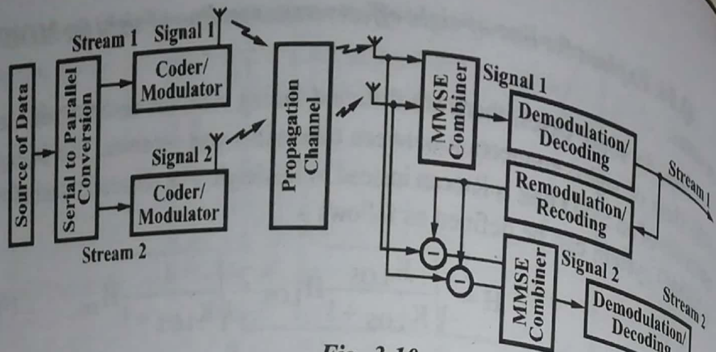


Fig. 3.10

(ii) Now, the required datastream is demodulated and decoded. Outputs from that process are firm decisions on the bits of stream 1. It should be noted that, we only require knowledge of the first stream to complete the decoding process, because we have separate encoding of distinct datastreams.

(iii) Now, these decoded bits are re-encoded and remodulated. We get the contribution that stream 1 has made to the total received signal at the distinct antenna elements by multiplying this symbol stream by the channel transfer function.

(iv) These contributions are subtracted from the signals at the distinct antenna elements.

(v) The second datastream can be detected by considering the cleaned-up signal. Here, we have N_r received signals, and $N_t - 2$ interferers. Now, the required datastream with diversity order 2 can be received using optimum combining again.

(vi) In the next step, the second datastream is decoded, recoded and remodulated. The associated signal is subtracted from the total signal at the receive antenna elements determined in the previous step. This further cleans-up the received signal.

The H-BLAST would be normal serial interference cancellation, when distinct transmit streams were to come from distinct users. It should be noted that the encoding scheme does not need *cooperation* between distinct users or antenna elements. When the first decoded stream is of bad quality, the H-BLAST faces the error propagation problem. It means that we subtract the wrong signal from the remaining signals at the antenna elements, when datastream 1 is not decoded correctly. Hence, even more interference is introduced instead of *cleaning-up* the receive signal. This further increase the probability of incorrect decoding of second datastream and so on. This problem can be eliminated by stream ordering. It means the receiver first decode the stream with best SINR, then the one with best next best, and so on.

Q.26. Write a short note on multiuser MIMO. ***(R.G.P.V., June 2016)***

Or

Discuss the working of multiuser MIMO. ***(R.G.P.V., June 2017)***

Ans. A multiuser MIMO system setup is shown in fig. 3.11. A signal base station with N_{BS} antenna elements shares information with K-mobile stations with $N_{MS}^{(k)}$ antenna elements each. Consider that N_{BS} is greater than $N_{MS}^{(k)}$ but that N_{BS} is lesser than $\sum N_{MS}^{(k)}$. Typically, a fourth generation cell would have $N_{BS} = 8$, $N_{MS}^{(k)} = 2$, and $K = 20$. There is no needed that all the mobile stations are active all the time. Instead of the base station can schedule the communication with specific users according to their state of channel. Therefore, certain fairness and delay criteria should be satisfied for the distinct users.

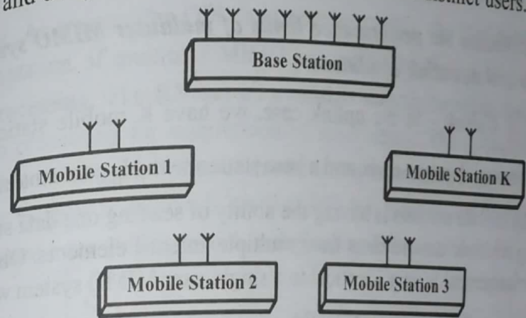


Fig. 3.11 Multiuser MIMO System

Q.27. Explain in detail the principle of multi input multi output system also explain multiuser MIMO. ***(R.G.P.V., June 2014)***

Ans. Refer to the ans. of Q.14 and Q.26.

Q.28. Write a detailed note on multiple input multiple output systems. What do you understand by multiuser MIMO ? ***(R.G.P.V., June 2015)***

Ans. Refer to the ans. of Q.16 and Q.26.

Q.29. What are the differences between single user MIMO and multiuser MIMO systems ?

Ans. Following are the main differences between single-user MIMO and multiuser MIMO systems –

(i) Cross-layer design is critical, and can largely decrease the breakdown probability of the overall capacity because of the channel state peculiarities. A judicious scheduling algorithm can largely decrease the possibility that the channels to two users are linearly dependent if there are more mobile stations as compared to there are base station antennas.

(ii) In the downlink, channel state information at the transmitter (CSIT) has a striking impact on the multiuser MIMO system performance. Hence, the followings conditions could take place without CSIT –

(a) The base station sends only $\min_k (N_{MS}^{(k)})$ streams, which drastically controls the downlink capacity.

(b) The base station sends N_{BS} data streams, but as the mobile stations cannot decode that several streams of data, the SINR is seriously worse at each mobile station, and hence the capacity is low.

Although, with CSIT, the base station can null out undesired streams in the different mobile stations direction, and hence send more data streams while still enabling at each mobile station a good SINR.

Q.30. Discuss the performance limits of multiuser MIMO system for the uplink and downlink conditions.

Ans. (i) Uplink – In the uplink case, we have K mobile stations with $(N_{MS}^{(k)})$ antenna elements each, and a base station with N_{BS} antenna elements. Each of the mobile stations is having the ability of sending one data stream or of sending multiple data streams from multiple antenna elements. Obviously, such an arrangement is very identical to a single-user MIMO system with N_{BS} receive antenna elements and $\sum_k N_{MS}^{(k)}$ transmit antenna elements. The only difference lies in the fact that the different mobile stations cannot help in sending their information. Although, this hardly affects the capacity.

Hence, the uplink system capacity increases with N_{BS}, the number of base station antenna elements.

(ii) Downlink – For the downlink case, the received signal $\mathbf{r}^{(k)}$ at the k-th mobile station can be defined as –

$$\mathbf{r}^{(k)} = \mathbf{n} + \mathbf{H}^{(k)}\tilde{\mathbf{s}} \qquad ...(i)$$

where $\tilde{\mathbf{s}}$ denotes the signal sent to the users. It comprises of a superposition of the data intended for all the distinct users $\tilde{\mathbf{s}} = \sum_k \tilde{\mathbf{s}}^{(k)}$. The $\tilde{\mathbf{s}}^{(k)}$ can be linearly or nonlinearly encoded versions of the signal streams intended for the distinct users. The mobile stations (RXs) cannot help for the datastreams decoding. This has much more important results for the practical implementation than the inability to help for the encoding i.e., access of the decoder to all the streams of data is critical for suppression of interference. As a result, the transmitter precoding is used to avoids interference at the mobile station.

When doing the encoding of multiple data streams, the order in which the streams are encoded is very important. The encoder for the k-th mobile station deals the interference from data streams for MS 1, ... k–1 as known, on the other hand the interference for users k + 1.....K has the same effect as noise. Hence, the first encoded stream has a worse SINR, on the other hand the later encoded streams have better SINR.

Q.31. Discuss the linear precoding method for the uplink.

Ans. At transmitter (TX) and receiver (RX), the simplest practical implementation of multiuser MIMO processing for the uplink is based on linear processing. The RX has all available signals, therefore can perform optimum processing for suppression of interference. The optimum transmit and receive weights depend on each other, thus an iterative computation of the optimum weights is required. We consider the minimization of the overall mean square error (MSE), which is defined as follows –

$$\text{MSE} = \text{tr}\left\{\sum_{k=1}^{K}\left\{\sum_{j=1}^{K}(\mathbf{H}^{(j)})^{\dagger}(\mathbf{T}^{(j)})^{\dagger}\mathbf{W}^{(k)}(\mathbf{W}^{(k)})^{\dagger}\mathbf{H}^{(j)}\mathbf{T}^{(j)} - (\mathbf{H}^{(k)})^{\dagger}(\mathbf{T}^{(k)})^{\dagger}\mathbf{W}^{(k)} - (\mathbf{W}^{(k)})^{\dagger}\mathbf{H}^{(k)}\mathbf{T}^{(k)} + \mathbf{I} + \sigma_n^2(\mathbf{W}^{(k)})^{\dagger}\mathbf{W}^{(k)}\right\}\right\} \qquad ...(i)$$

where $(W^{(k)})^{\dagger}$ denotes the receive matrix for the k-th user. Then, the aim is to decrease this mean square error under the power constraints as follows –

$$\text{tr}\left\{(\mathbf{T}^{(k)})^{\dagger}\mathbf{T}^{(k)}\right\} \le P_k^{max} \qquad ...(ii)$$

The optimum RX weights are functions of the transmit weight of all users, on the other hand the TX weights depend on the receive weights of all users. Hence, the receive and transmit weights are obtained as –

(i) Update for all users (k = 1, 2,.... K)

$$(\mathbf{W}^{(k)})^{\dagger} = (\mathbf{T}^{(k)})^{\dagger}(\mathbf{H}^{(k)})^{\dagger}\left[\sigma_n^2\mathbf{I} + \sum_{j=1}^{K}\mathbf{H}^{(j)}\mathbf{T}^{(j)}(\mathbf{H}^{(j)})^{\dagger}(\mathbf{T}^{(j)})^{\dagger}\right]^{-1} \qquad ...(iii)$$

(ii) Update for all users (k = 1, 2, ..., K)

$$\mathbf{X}^{(k)}(\mu'_k) = \left[\mu'_k\mathbf{I} + \sum_{j=1}^{K}\mathbf{H}^{(k)}\mathbf{W}^{(j)}(\mathbf{H}^{(k)})^{\dagger}(\mathbf{W}^{(j)})^{\dagger}\right]^{-1}(\mathbf{H}^{(k)})^{\dagger}\mathbf{W}^{(k)} \quad ...(iv)$$

$$\mu_k = \text{Max}\left[\arg_{\mu'_k}\left(\text{tr}\left\{\mathbf{X}^{(k)}(\mu'_k)(\mathbf{X}^{(k)}(\mu'_k))^{\dagger}\right\} = P_k^{max}\right), 0\right] \quad ...(v)$$

$$\mathbf{T}^{(k)} = \left[\mu_K\mathbf{I} + \sum_{j=1}^{K}\mathbf{H}^{(k)}\mathbf{W}^{(j)}(\mathbf{H}^{(k)})^{\dagger}(\mathbf{W}^{(j)})^{\dagger}\right]^{-1}(\mathbf{H}^{(k)})^{\dagger}\mathbf{W}^{(k)} \quad ...(vi)$$

In the typical case, 10 to 20 iterations are enough for convergence.

Q.32. Discuss the linear precoding method for the downlink.

Ans. In linear precoding (beamforming) at the TX, the total transmit signal intended for the k^{th} user as follows –

$$\tilde{\mathbf{s}}^{(k)} = \mathbf{T}^{(k)}\mathbf{s}^{(k)} \quad ...(i)$$

where the $\mathbf{T}^{(k)}$ denotes the precoding matrix for the k-th user. Since we suppose that the modulation symbols have unit energy $\mathbf{R}_{ss}^{(k)} = \mathbf{I}$, the correlation matrix of the k-th user's signal represents $\mathbf{R}_{\tilde{s}\tilde{s}}^{(k)} = E\left\{\tilde{\mathbf{s}}^{(k)}\tilde{\mathbf{s}}^{(k)\dagger}\right\}$, and the power assigned to the k-th user represents $P_k = \text{tr}\{\mathbf{R}_{\tilde{s}\tilde{s}}^{(k)}\} = \text{Trace}\{\mathbf{T}^{(k)}\mathbf{T}^{(k)\dagger}\}$.

(i) Block Diagonalizations Method – Let us assume that the case in which each mobile station (MS) has multiple antenna elements ($N_r^{(k)} \geq 1$) and the base station (BS) sends $N_r^{(k)}$ data streams to each user. A simple solution is obtained by imposing the dimensionality constraint $\sum_k N_r^{(k)} = N_t$.

According to this constraint, adding receive antennas at the mobile station reduces the number of users which is sent to at a given point in time. Clearly, this is not a reasonable constraint, as additional receive antennas perform to improve the reception quality. More precisely, it is the number of data streams that intended for the mobile stations which is limited, and data streams intended for distinct mobile stations (MSs) have

to be kept apart through suitable beamforming at the base station side by a block diagonalization method. An interference channel matrix $\tilde{\mathbf{H}}^{(k)}$ is defined for each user as follows –

$$\tilde{\mathbf{H}}^{(k)} = \left[(\mathbf{H}^{(1)})^T, ...(\mathbf{H}^{(k-1)})^T(\mathbf{H}^{(k+1)})^T, ...(\mathbf{H}^{(K)})^T\right]^T \quad ...(ii)$$

Any precoding matrix $\mathbf{T}^{(k)}$ that lies in the nullspace of $\tilde{\mathbf{H}}^{(k)}$ causes to a situation in which data streams intended for the other mobile stations do not affect the k-th mobile station (MS). As a result, there is now a new dimensionality constraint. Assume that J_k be the rank of $\tilde{\mathbf{H}}^{(k)}$. Then the block diagonalization is defined, as

$$N_t > \underset{k}{\text{Max}}(J_1, J_2, ..., J_K) \quad ...(iii)$$

When the propagation channels are all full rank, this implies that the number of transmit antennas should be no smaller as compared to the number of data streams. Then, the singular value decomposition of $\tilde{\mathbf{H}}^{(k)}$ can be defined as

$$\tilde{\mathbf{H}}^{(k)} = \tilde{\mathbf{U}}^{(k)}\tilde{\Sigma}^{(k)}[\tilde{\mathbf{V}}_{fc}^{(k)}\ \tilde{\mathbf{V}}_{lc}^{(k)}]^{\dagger} \quad ....(iv)$$

where $\tilde{\mathbf{V}}_{fc}^{(k)}$ comprises the first J_k and $\tilde{\mathbf{V}}_{lc}^{(k)}$ the last $N_t - J_k$ right singular vectors. Hence $\tilde{\mathbf{V}}_{lc}^{(k)}$ forms an orthonormal basis of the nullspace of $\tilde{\mathbf{H}}^{(k)}$. A new overall effective channel matrix can be defined as –

$$\hat{\mathbf{H}} = \begin{bmatrix} \mathbf{H}^{(1)}\tilde{\mathbf{V}}_{lc}^{(1)} & 0 & 0 \\ 0 & ..... & 0 \\ 0 & 0 & \mathbf{H}^{(K)}\tilde{\mathbf{V}}_{lc}^{(K)} \end{bmatrix} \quad ...(v)$$

Next the singular value decomposition of this overall effective matrix is determined. This is obtained in an efficient way because of the block-diagonal structure of $\hat{\mathbf{H}}$. For the k-th block,

$$\hat{\mathbf{H}}^{(k)} = \hat{\mathbf{U}}^{(k)}\begin{bmatrix}\hat{\Sigma}^{(k)} & 0 \\ 0 & 0\end{bmatrix}\left[\hat{\mathbf{V}}_{fc}^{(k)}\hat{\mathbf{V}}_{lc}^{(k)}\right]^{\dagger} \quad ...(vi)$$

Then, the complete beamforming/power allocation matrix **T** can be defined as –

$$\mathbf{T} = \left[\tilde{\mathbf{V}}_{lc}^{(1)}\hat{\mathbf{V}}_{fc}^{(1)}\ \tilde{\mathbf{V}}_{fc}^{(2)}\hat{\mathbf{V}}_{fc}^{(2)} .....\tilde{\mathbf{V}}_{lc}^{(K)}\hat{\mathbf{V}}_{fc}^{(K)}\right]\Lambda^{1/2} \quad ...(vii)$$

where Λ denotes the diagonal matrix. This parameters is used to perform the water filling on the elements of

$$\begin{bmatrix} \hat{\Sigma}^{(1)} & & & 0 \\ & \hat{\Sigma}^{(2)} & & \\ & & \ldots\ldots & \\ 0 & & & \hat{\Sigma}^{(K)} \end{bmatrix} \quad \text{...(viii)}$$

(ii) Joint Wiener Filtering Method – The mean square error can be determined as follows –

$$MSE_k = E[\|\mathbf{r}_k - \mathbf{z}_k\|^2] \quad \text{...(ix)}$$

where the expectation is over the random data vectors, $\{\mathbf{s}_k\}_{k=1}^{K}$, and the noise, $\{\mathbf{n}_k\}_{k=1}^{K}$. Then, the optimization problem can be defined as

$$\left\{\mathbf{T}_k^{opt}\right\}_{k=1}^{K} = \arg \min_{[\mathbf{T}_k]_{k=1}^{K}} \sum_k MSE_k \ \text{ for } k = 1, \ldots.. K \quad \text{...(x)}$$

$$\text{s.t. } \mathrm{Tr}\left\{\sum_k \mathbf{T}_k^{\dagger}\mathbf{T}_k\right\} \le P_{max}$$

This optimization problem is solved by taking an eigendecomposition $\mathbf{V}\Lambda\mathbf{V}^{\dagger}$ as –

$$\mathbf{H}_k^{\dagger}\left[\sigma_n^2\mathbf{I} + \sum_{j\neq k}\mathbf{H}_i\,\mathbf{T}_j\,\mathbf{H}_j^{\dagger}\,\mathbf{T}_j^{\dagger}\right]\mathbf{H}_k = [\mathbf{V}_k\,\overline{\mathbf{V}}_k]\begin{bmatrix}\Lambda_k & \\ & \overline{\Lambda}_k\end{bmatrix}[\mathbf{V}_k\,\overline{\mathbf{V}}_k]^{\dagger} \quad \text{....(xi)}$$

where $\mathbf{V}_k$ denotes a square matrix of dimension rank $\mathbf{H}_k$. Then the optimum precoding matrix can be defined as –

$$\mathbf{T}_k = \mathbf{V}_k\begin{bmatrix} 0 & 0 & \ldots.. & 0 & \xi_{1,k} & 0 & 0 \\ \ldots.. & \ldots.. & \ldots.. & \ldots.. & 0 & \ldots.. & 0 \\ 0 & \ldots.. & \ldots.. & 0 & 0 & 0 & \xi_{L,k} \end{bmatrix} \quad \text{....(xii)}$$

where L denotes the number of spatial streams and

$$\xi_{i,k} = \mathrm{Max}\left[\frac{1}{\mu^{1/2}[\Lambda_k]_{i,i}^{1/2}} - \frac{1}{[\Lambda_k]_{i,i}}, 0\right] \quad \text{....(xiii)}$$

and μ is selected to fullfill the total power constraint.

The optimum RX filter matrix represents the Wiener filter for a given transmit precoder as –

$$\mathbf{W}_k = \left[\mathbf{H}_k\mathbf{T}_k\mathbf{H}_k^{\dagger}\mathbf{T}_k^{\dagger} + \left[\sigma_n^2\mathbf{I} + \sum_{j\neq k}\mathbf{H}_i\mathbf{T}_j\mathbf{H}_j^{\dagger}\mathbf{T}_j^{\dagger}\right]\right]^{-1}\mathbf{H}_k\mathbf{T}_k \quad \text{....(xiv)}$$

(iii) Joint Leakage Suppression Method – In this method, the precoding matrices can be used to increase the signal-to-leakage and noise-ratio (SLNR). The SLNR of the k-th user can be defined as –

$$SLNR_k = \frac{E\{\mathbf{s}_k\mathbf{H}_k\mathbf{T}_k\mathbf{H}_k^{\dagger}\mathbf{T}_k^{\dagger}\mathbf{s}_k\}}{N_r\sigma_n^2 + E\{\Sigma_{i\neq k}\Sigma_{j\neq k}\mathbf{s}_k\mathbf{H}_j\mathbf{T}_k\mathbf{H}_i^{\dagger}\mathbf{T}_k^{\dagger}\mathbf{s}_k\}} \quad \text{....(xv)}$$

or

$$SLNR_k = \frac{\mathrm{Tr}[\mathbf{H}_k\mathbf{T}_k\mathbf{H}_k^{\dagger}\mathbf{T}_k^{\dagger}]}{\mathrm{Tr}\left[\mathbf{T}_k^{\dagger}\left[N_r\sigma_n^2\mathbf{I} + \tilde{\mathbf{H}}_k^{\dagger}\tilde{\mathbf{H}}_k\right]\mathbf{T}_k\right]} \quad \text{....(xvi)}$$

Hence, the base station requires to determine precoding matrices $\mathbf{T}_k$ which increase the $SLNR_k$ subject to a transmit power constraint $\mathrm{Tr}\{\mathbf{T}_k^{\dagger}\mathbf{T}_k\} = P_{max}$. We can obtain an auxiliary matrix **A** as a matrix which satisfies

$$\mathbf{A}_k^{\dagger}\mathbf{H}_k^{\dagger}\mathbf{H}_k\mathbf{A}_k = \Lambda_k$$

$$\mathbf{A}_k^{\dagger}\left[N_r\sigma_n^2\mathbf{I} + \tilde{\mathbf{H}}_k^{\dagger}\tilde{\mathbf{H}}_k\right]\mathbf{A}_k = \mathbf{I} \quad \text{....(xvii)}$$

where Λ_k denotes a diagonal matrix with nonnegative entries, sorted in descending order. Then, the precoding matrix **T** can be the first L_k columns of **A**.

If each user has only one data stream, L_k is equal to one, the beamforming matrix becomes a vector, and is achieved in a easy manner i.e., perform an eigendecomposition of the matrix.

$$[N_r\sigma_n^2\mathbf{I} + \tilde{\mathbf{H}}_k^{\dagger}\tilde{\mathbf{H}}_k]^{-1}\mathbf{H}_k^{\dagger}\mathbf{H}_k$$

The optimum precoding vector denotes the eigenvector assigned with the highest eigenvalue of this matrix.

Q.33. What do you mean by closed-loop systems and quantized feedback ?

Ans. A better way of decreasing the feedback overhead for send beamforming is provided by limited-feedback codebooks. Thus, the technique is very complex for multiuser MIMO. First of all, the quantization of the send precoding vector has to be very finer than in the single-user case. In a single user case the base station makes a beam pattern which represents a maximum in the targeted mobile station direction. Because of the quantization influences slight deviations from the optimum beam pattern do not cause a significant loss of signal-to-noise-ratio (SNR). In the multiuser case, we require to guarantee low interference by signals from other users to achieve good SINR at the mobile stations. Alternatively, each send beam pattern requires to place nulls toward the users it should not cover. Now, we know that nulls are very sensitive to perturbations of the antenna weights than "main beams". Hence, the quantization effects in multiuser settings should be finer.

Another complication originates from the fact that the optimum settings of the precoder depend on the all users channels. However, the mobile station cannot easily determine the precoding matrix setting that it requires to optimize its performance. This problem can be solved by projection techniques. Firstly, here the base station transmits out a number of training signals on Q = K distinct beams $g^{(q)}$. Then, the k-th mobile station forms the quantities as follows –

$$\gamma_k = \underset{q}{\mathrm{Max}} \frac{|\mathbf{h}^{(k)}(\mathbf{g}^{(q)})^T|^2}{\sigma_n^2 + \sum_{p \neq q} |\mathbf{h}^{(k)}(\mathbf{g}^{(p)})^T|^2}$$

Simply, which is the maximum SINR it can obtain if the base station sends to this mobile station on the beam which is optimum for it, and to all the other mobile stations on distinct beams, without any control of power. Then, the k-th mobile station just feeds back the index q with which it obtaines this SINR.

In the spirit of the random beamforming approach, it is advantageous to change the beams onto which the projection occur. Alternatively, the $g^{(q)}$ is randomly produced. Then, the probability is more evenly distributed for each mobile station to observe a channel which increases its SINR and hence increase its throughput.

Q.34. *Write short note on base station cooperation.*

Ans. In a cellular system, the existence of multiple users, and thus interferers, influences the capacity, and decreases the data rate which is possible for a single user. Base station cooperation eliminates the multiuser interference.

For the uplink, cooperating base stations are observed as a giant MIMO system with $N_r N_{BS}$ antenna elements at one link end, in which N_{BS} denotes the number of cooperating base stations.

For the downlink, suitable preprocessing reduces the effect of interfering base stations, if the base stations cooperate and each base station has the

channel state information for all mobile stations. One basic distinction of multi-BS systems originates from the fact that the interference reaching at the undesired mobile stations is basically asynchronous. Considering perfect timing synchronization among cooperative base stations, the timing-advance process can ensure that the desired signals for an mobile station which are sent from multiple base stations arrive the mobile station at the same time. Therefore, the undesired signals do not reach anymore synchronously. This causes to a change in the interference statistics, and needs modifications of how to determine the precoding coefficients. We have multiple power constraints, one for each base station, this is the another main difference to multiuser MIMO. This complicates the design of precoder.

In the literature, the base station cooperation is called network MIMO or cooperative multi-point (CoMP). It is intended for use in fourth-generation cellular systems, because it can greatly improve performance at the cell edge particularly.

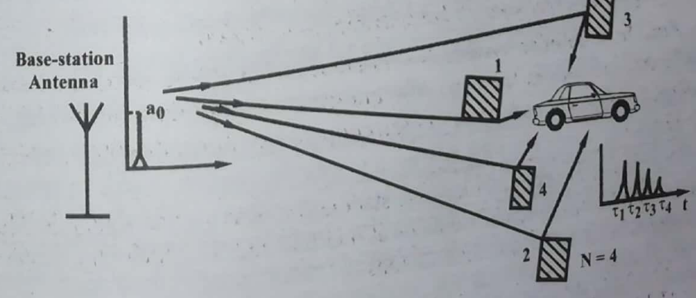
Q.35. *What is delay spread ?* **(R.G.P.V., Dec. 2016)**

Ans. If the base station (BS) transmits an impulse signal $f_0(t) = a_0\delta(t)$ to the mobile unit, then the phenomenon of delay spread takes place. The impulse signal becomes many of impulse signals which arrive at the mobile unit at different times when due to multipath scattering. The total delay spread time is approximately lengthened and the phenomenon is such as the voice echoes received on the top of the mountain. Fig. 3.12 (a) shows the delay spread for 4-scatterer case and fig. 3.12 (b) shows the delay spread for N-scatterer. At the mobile unit, the received impulse signal can be expressed as –

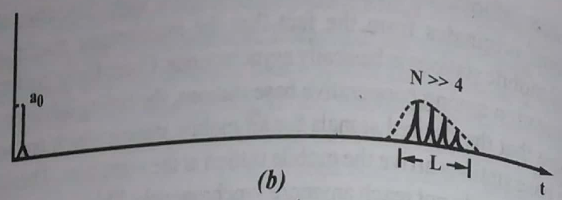
$$f(t) = a_0 \sum_{i=1}^{n} a_i \delta(t - \tau_i).e^{j\omega t} = E(t)e^{j\omega t} \quad ...(i)$$

Here,

$$E(t) = a_0 \sum_{i=1}^{n} a_i\, \delta(t - \tau_i)$$



(a)



(b)

Fig. 3.12

The above equation denotes a train of discrete impulses at frequency ω arriving at the mobile receiver [fig. 3.12 (a)]. The received discrete impulses become a continuous signal pulse with a length of pulse L if the several scatterers in the vicinity of the mobile unit increases, called the delay spread [see fig. 3.12 (b)].

OO

UNIT 4

COGNITIVE RADIO

PROBLEM DESCRIPTION, COGNITIVE TRANSCEIVER ARCHITECTURE, PRINCIPLE OF INTERWEAVING, SPECTRUM SENSING

Q.1. Define cognitive radio.

Ans. All transmission parameters to the environment, like bandwidth, coding, modulation format, transmission times etc. are used by a fully cognitive radio. A fully cognitive radio is currently complicated for practical purposes, but from a scientific point of view it is interesting.

Depending upon the environment, a spectrum sensing cognitive radio only employs bandwidth, transmission frequency and time. This radio is also known as dynamic spectrum access (DSA).

Q.2. Write down some applications of cognitive radio. *(R.G.P.V., June 2016)*

Ans. Some applications of cognitive radio are given below –

(i) Opportunistic cognitive radio networking.

(ii) T.V. white spaces.

(iii) Cognitive radio for emergency and public saftey applications.

(iv) Operator-controlled/assisted cognitive radio networks.

(v) Novel applications for cognitive radio network.

Q.3. Define self aware cognitive radio. *(R.G.P.V., Dec. 2016)*

Ans. Self-awareness refers to the ability of the radio to understand its own capabilities, i.e., to understand what it does and does not know, as well as the limits of its capabilities. In this way, the radio can obtain whether a task is within its capabilities. In the case of a basic self-aware radio, it should know its current performance, such as bit error rate (BER), signal-to-interference and noise ratio (SINR), multipath and others. For instance, for the radio to assess its travel speed a fortnight ago between locations A and B, it might be able to extract parameters from its log file and do the calculation. For the radio to decide whether it should search for the specific entries in the log and then perform appropriate calculations, it requires to know the effort needed to perform such a task and the needed accuracy of the estimate to its current task.

Q.4. List and explain the characteristic of radio cognition task.
(R.G.P.V., Dec. 2016)

Ans. Cognition radio network features which enable more efficient and flexible usage of the spectrum are given below –

(i) Dynamic Frequency Selection – It is the ability to dynamically choose the optimal frequency from the sensed spectrum.

(ii) Transmit Power Control – The transmission power must be adaptive to improve spectrum utilization.

(iii) Frequency Ability – It is the ability to adapt to varying operating frequency environment.

(iv) Adaptive Modulation – It is the reconfiguration of transmission characteristics and waveforms to enable efficient spectrum usage.

Q.5. Give the classification of models for dynamic spectrum access.

Ans. Classification of models for dynamic spectrum access are as follows –

(i) Hierarchical Access Model – Hierarchical access model allocates different priorities to different users. Primary customers would be handled in such a way that they experience the similar service quality when the spectrum were reserved exclusively for them. Secondary customers are permitted to send, however only in such a way that they don't affect the performance of the primary customers. The secondary customer adaptively chooses whether they might employ parts of a spectrum which is given by default to primary customers. In other words, spectrum which is given to a primary customer, but it is temporarily available is exploited by a cognitive radio.

(ii) Dynamic Exclusive Model – In this model, various providers can share the spectrum, however, a frequency band is reserved for the exclusive use of a certain service. Combined usage of a single high frequency band results in higher spectral efficiency than the use of N smaller bands by N separate operators because of trunking gain. Trading can also be used for spectrum sharing.

(iii) Open Sharing Model – Depending upon some constraints on the characteristics of the transmitted signal, all customers can access the spectrum equally. At present this technique is used for industrial, scientific and medical (ISM) bands. Merits and demerits of this technique can be seen

when wireless fidelity devices are proliferated in this band. However, it has become impossible for other services to operate in this frequency band with good quality of service, due to many wireless fidelity devices.

Q.6. Give the reasons, why primary users would permit secondary users to employ their spectrum.

Ans. Primary users would permit secondary users to employ their spectrum due to the following reasons –

(i) Regulatory Requirements – The frequency regulator may instruct that cognitive devices can use a fix range of spectrum as long as they do not interfere with primary customers. This type of technique is likely for parts of the spectrum that primary customers never paid for. In U.S.A., television stations did not purchase the spectrum they use but rather got it for free since they are deemed to perform a public service. This makes it simple for the frequency regulator to demand that television stations coexist with other services in public interest.

(ii) Profit – Owners of the spectrum can charge secondary customers. Traditionally auction systems have been adopted in which secondary customers could purchase in real time the right to a particular part of spectrum for a short time. This is promising for few uses, but cannot be used always because the costs for billing and monitoring would become greater than the revenue generated from the auctioning of the spectrum.

(iii) Emergency Services – In emergency time, another form of cognitive radio appears, when primary customers have to provide spectrum for emergency services.

According to the principle of hierarchical cognitive radio, secondary customers do not disturb the primary customers. This nondisturbance may be obtained by interweaving, underlay and overlay techniques.

Q.7. Explain about the cognitive transceiver architecture.
(R.G.P.V., June 2014)

Or

Explain cognitive transceiver architecture in detail.
(R.G.P.V., June 2015)

Or

Explain cognitive radio with the help of transceiver architecture
(R.G.P.V., June 2016)

Or

Discuss the cognitive transceiver architecture. ***(R.G.P.V., June 2017)***

Or

Describe the architecture of the cognitive transceiver.
(R.G.P.V., May 2018)

Ans. The block diagram of a cognitive transceiver architecture is shown in fig. 4.1. Analog to digital converter, radio frequency and baseband processing are the three important parts of cognitive transceiver architecture.

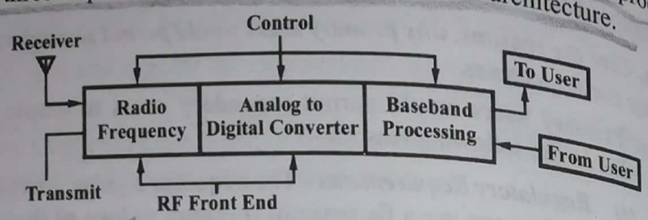


Fig. 4.1 Structure of a Cognitive Transceiver

One or more of those parts is made adaptive depending on the specific type of cognitive radio. The radio frequency front end is different from a conventional receiver, in a spectrum sensing cognitive radio. The block diagram of radio frequency receiver front end is shown in fig. 4.2.

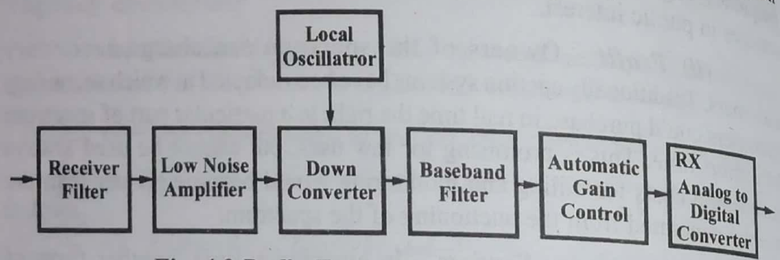


Fig. 4.2 Radio Frequency Receiver Front End

All the components, such as low noise amplifier, antennas, AGC, local oscillator of radio frequency receiver front end have to be wideband so that they can operate at all possible frequencies at which the cognitive radio might wish to operate. By using an adjustable local oscillator the front end can choose the channel which the cognitive radio is to use. A major challenge in cognitive radio lies in the possibility of receiver saturation via strong out-of-band signals because the channel choice happens only after the mixer (downconverter). A signal which is not in the band that the receiver wishes to demodulate at a particular time, however still within the overall receive band of cognitive radio can saturate the radio frequency component like low noise amplifier. Radio frequency notch filters can reduce strong interference, when placed before low noise amplifier. However, the tunablity of such filters is limited and the hardware cost is high. The baseband processing also has to be adaptive for a fully cognitive radio. This can be simply obtained by implementing the baseband processing as software on a digital signal processor.

Q.8. Draw cognitive radio framework and explain each block.
(R.G.P.V., Dec. 2016)

Ans. Refer to the ans. of Q.7.

Q.9. What is cognitive cycle ?
(R.G.P.V., Dec. 2016)

Ans. The main cognitive tasks needed to obtain dynamic sharing are shown in fig. 4.3, and can be known as cognitive cycle. The three main steps of this cycle are given below –

(i) Spectrum Sensing – It is defined a monitoring the available spectrum bands to detect spectrum holes. The main challenge is to do this energy efficiently. In addition, the cost of hardware for this spectrum sensing must be considered into account. Next to sensing the spectrum, it is necessary to monitor other information in the wireless communication scene.

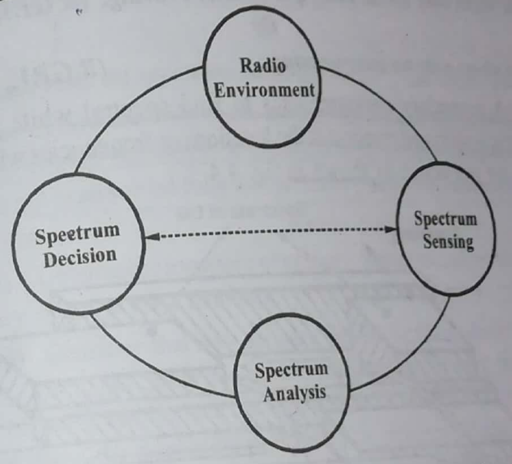


Fig. 4.3 Cognitive Cycle for Dynamic Sharing

(ii) Spectrum Analysis – It is useful to build a model of the wireless communication scene based on the measurements obtained by spectrum sensing. Because of hardware and energy budget limitations, it will be impossible to monitor the whole spectrum or scene continuously and in great detail. Also, wireless networks are spread in space which makes it difficult to build a model of the full spectrum scene. To improve the model accuracy, cooperation between nodes has been introduced, at the cost of increased communication overhead. Frequently, the spectrum model will be built on local partial information. Methods for local spectrum analysis are an important research objective for this task.

(iii) Spectrum Decision – It is about whether and how to access the spectrum. The more reconfigurability present in the cognitive radio, the more optimization options are available. The optimal spectrum access option is the one that maximizes the application or user requirements given the environment or spectrum policy constraints. The spectrum decision is framed as an optimization problem using the model built during spectrum analysis. This optimization problem may have a local or a global optimization goal. Next the spectrum decision must be communicated across users, which is often done through a common channel. Because the availability of such channel cannot be relied on in the context of dynamic opportunistic spectrum access, the major spectrum decision challenge relates to the development of spectrum coordination methods that do not rely on such a common channel.

Q.10. *Discuss the principle of interweaving in cognitive radio.* **(R.G.P.V., June 2014)**

Or

Write short note on the principle of interweaving. **(R.G.P.V., June 2015)**

Or

Write short note on interweaving **(R.G.P.V., June 2016)**

Ans. A secondary customer tries to find spectral white space in an interweaving system and transmit at the locations or frequencies where primary customers are not active as shown in fig. 4.4.

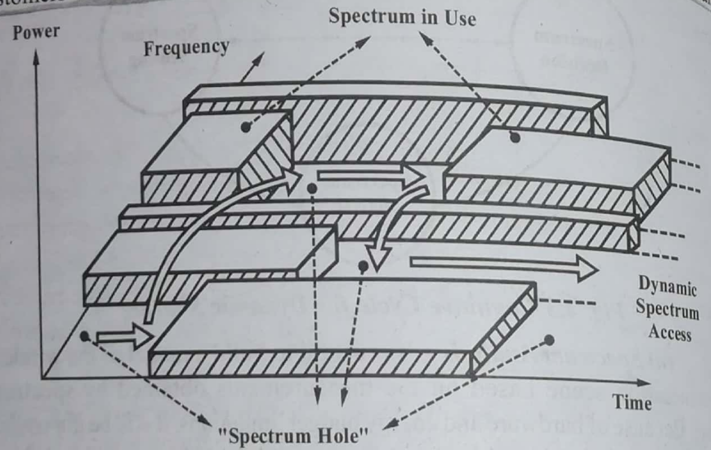


Fig. 4.4 Concept of Spectrum Holes

This strategy consists following three steps as follows –

(i) Spectrum Sensing – The parts of the time frequency plane which are not used by primary customers have to be identified by the cognitive radio. In the presence of noise, the sensing cannot be fully reliable. Further, it is essential to find the signal level of the detected radiation. Spectrum sensing provides essential information for spectrum sharing. Interference at a particular receiver is mostly characterized by the interference temperature which explains the interfering PSD divided by Boltzmann's constant k_B.

(ii) Spectrum Management – In spectrum management, secondary system decides when and in which part of the spectrum, to transmit. Because of many reasons this decision is difficult, some of them are given below –

(a) Secondary system decision is based on the sensing information which is not perfect.

(b) Secondary system has only causal knowledge of the spectrum occupation.

(iii) Spectrum Sharing – Free spectrum division decision is known as spectrum sharing. Spectrum management and spectrum sharing are strongly related because a particular part of spectrum can be useful for a specific secondary transmitter while not for another.

Q.11. *Explain spectrum sensing in a hierarchical system.*

Or

Write short note on spectrum sensing. **(R.G.P.V., June 2016)**

Ans. Any device can only sense transmitted radiation is one of the main problem of spectrum sensing in a hierarchical cognitive radio system. This situation has not been solved fully though many possible solutions are available as given below –

(i) Geographic locations of primary receivers can be made available to secondary systems via a database.

(ii) Consider that any primary device works as both receiver and transmitter. Then, the presence of the primary device may be determined from the radiation transmitted by that device. A secondary system requires knowledge about the duplexing policies of the primary device. For instance, when a frequency division duplexing primary transmits in a specific band, the secondary system would know in which band such a primary device would receive.

(iii) A common control channel which permits dissemination of information about spectrum usage. This method permits greatly efficient utilization of the spectrum, however has the disadvantage that existing devices require to be modified to permit signaling on this channel.

(iv) The primary receiver's spurious emissions are observed. For instance, the emissions of the local oscillators of television receivers can be observed by the secondary system because their local oscillators are not perfectly isolated from the antenna. This technique has the disadvantage that the level of the observed signal cannot simply be mapped to the location of the receiver and that it really punishes good receiver design.

(v) Let, primary receivers be in all possible locations where primary transmissions may be heard at a level sufficient for demodulation. This type of policy is very conservative. Receivers which need to be protected are considered to be everywhere in the coverage area of the primary system.

Q.12. Give the classification of detectors.

Ans. There are three types of detectors employed for sensing, based on the amount of a priori knowledge that sensor has about the transmitted waveform –

(i) Energy detection (ii) Matched filter (iii) Cyclostationarities.

(i) Energy Detection – If the sensor has no information about the signal structure energy detection is employed. If the sensor does not know about the waveform sent by other customers in a frequency band of interest, it may only determine the energy available in that band. The sensor determines only thermal noise energy if no other customer is available, otherwise sensor determines signal plus noise energy. Signal is available or not is a classical detection problem which can be formulated as hypothesis testing. Assume a single narrowband channel for simplicity, where the received signal is

$$r_n = n_n \quad \mathcal{H}_0 : \text{noise only received} \quad ...(i)$$

$$r_n = hs_n + n_n \quad \mathcal{H}_1 : \text{noise plus signal received} \quad ...(ii)$$

where r_n = Received signal at time n

s_n = Transmitted signal at time n

h = Channel gain

n_n = Observed noise.

$\mathcal{H}_0$ and $\mathcal{H}_1$ = Hypotheses between the sensor requires to decide.

The obtained signal samples may be averaged over a number time instances N, hence the decision variables are used given below –

$$y = \sum_{n=1}^{N} |r_n|^2 \quad ...(iii)$$

When the number of samples is large, y are Gaussian random variables with expected value and variance.

$$E\{y\} = \begin{cases} N\sigma_n^2 & \mathcal{H}_0 \\ N[|h|^2 + \sigma_n^2] & \mathcal{H}_1 \end{cases} \quad ...(iv)$$

$$\sigma_y^2 = \begin{cases} 2N\sigma_n^4 & \mathcal{H}_0 \\ 2N\sigma_n^2[2|h|^2 + \sigma_n^2] & \mathcal{H}_1 \end{cases} \quad ...(v)$$

Then the decision rule is

$$y \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \theta \quad ...(vi)$$

where θ = Threshold.

The threshold cannot be set to perfectly detect the presence of a signal, because of the random noise. The probability of a false alarm is as follows –

$$P_f(\theta) = \Pr(Y > \theta \mid \mathcal{H}_0)$$

$$= Q\left(\frac{\theta - N\sigma_n^2}{\sigma_n^2\sqrt{2N}}\right) \quad ...(vii)$$

Similarly, the probability for the missed detections is

$$P_{md}(\theta) = \Pr(Y < \theta \mid \mathcal{H}_1)$$

$$= 1 - Q\left(\frac{\theta - N[|h|^2 + \sigma_n^2]}{\sigma_n\sqrt{2N[2|h|^2 + \sigma_n^2]}}\right) \quad ...(viii)$$

We can trade off between false-alarm and missed-detection probability by adjusting the threshold. Generally, the frequency regulator determines the missed-detection probability in hierarchical cognitive system because a missed detection of the spectrum sensor implies that the secondary system will transmit even though a primary customer is active in the assumed band.

(ii) Matched Filter – If the sensor knows the transmitted waveform, the spectrum sensing performance may be improved considerably. In such conditions the best way of improving the signal-to-noise ratio of the detection process is matched filtering. The output of the matched filter, for a spectrum sensing cognitive radio, is employed as a test statistics to detect whether signal energy is available in the assumed band or not. The symbols transmitted by the primary user have to detect by a fully cognitive radio. In some situations, a sensor might wish to decode only beacons since those have generally better signal to noise ratio compared to the real payload data, and might also have information about spectrum usage which can be used by a secondary system.

(iii) Cyclostationarities – Transmitted signals generally contain cyclostationary statistics i.e., some statistical properties are periodic. This fact may be used to enhance the performance of an energy detector. Hence, the incoming signal is correlated with itself before being transmitted to an energy detector.

Q.13. Explain in detail multinode detection.

Ans. The secondary system have multiple nodes that can help each other to reach a better sensing decision in many cases. Every secondary node listens and makes an independent decision concerning whether it may sense a fixed channel to be occupied. This binary information is exchanged by the nodes to reach a combined decision as to whether the spectrum is present or not. The secondary system should decide that a frequency band is taken if at least one node senses occupation to protect the primary system.

Hence, the overall false alarm probability increases as follows –

$$P_{f,\ (network)} = 1 - \prod_{k=1}^{K}[1 - P_{f,k}] \qquad ...(i)$$

and the probability of missed detection will decrease

$$P_{md,\ network} = \prod_{k=1}^{K} P_{md,k} \qquad ...(ii)$$

If the nodes instead of communicating just a binary decision about spectrum occupancy, perform a linear combination of the determined averaged samples y, a more precise joint decision can be made. For the signal from the k-th node introducing linear weights w_k, we can form a total decision variable,

$$z = \sum w_k y_k = \mathbf{w}^T\mathbf{y} \qquad ...(iii)$$

The decision variable is a weighted sum of Gaussians. Hence, its mean and variance are –

$$E\{z\} = \begin{cases} N\sigma_n^2\mathbf{w}^T\mathbf{1} & \mathcal{H}_0 \\ N\mathbf{w}^T[\mathbf{g} + \sigma_n^2\mathbf{1}] & \mathcal{H}_1 \end{cases}$$

$$\sigma_z^2 = \begin{cases} 2N\sigma_n^4\mathbf{w}^T\mathbf{w} & \mathcal{H}_0 \\ 2N\sigma_n^2\mathbf{w}^T[2\ \mathrm{diag}(\mathbf{g}) + \sigma_n^2\mathbf{I}]\mathbf{w} & \mathcal{H}_1 \end{cases}$$

Here, $\mathbf{g} = [|h_1|^2, |h_2|^2, \ldots\ldots |h_k|^2]^T$

Diag(**g**) = Diagonal matrix whose entries are the $|h_k|^2$.

Then, for false alarm and missed detection, probabilities are –

$$P_f(\theta, \mathbf{w}) = Q\left(\frac{\theta - N\sigma_n^2\mathbf{w}^T\mathbf{1}}{\sigma_n^2\sqrt{2N\mathbf{w}^T\mathbf{w}}}\right)$$

$$P_{md}(\theta, \mathbf{w}) = 1 - Q\left(\frac{\theta - N\mathbf{w}^T[\mathbf{g} + \sigma_n^2\mathbf{1}\mathbf{1}]}{\sigma_n\sqrt{2N\sigma_n^2\mathbf{w}^T[2\ \mathrm{diag}(\mathbf{g}) + \sigma_n^2\mathbf{I}]\mathbf{w}}}\right)$$

Now, we have to define the maximum weights and thresholds for the detection. It is clear that communicating the full decision variables from every sensing node needs more overhead than just communicating a binary decision. In contrast it is rarely possible to really just send a single bit over a wireless channel.

Q.14. *Explain the cognitive pilots in brief.*

Ans. Primary customers sensing can be highly facilitated by the cognitive pilot channels. The following phases are involved in the use of CPC –

(i) At the initialization, the wireless terminal first listen to the cognitive pilot channels.

(ii) The wireless network obtains the information and chooses the most useful one to setup its communication.

(iii) The cognitive pilot channel is broadcasted to a wide area.

SPECTRUM MANAGEMENT, SPECTRUM SHARING, OVERLAY, UNDERLAY

Q.15. *Explain the spectrum opportunity tracking.*

Or

Write short notes on the spectrum management.

(R.G.P.V., June 2015, 2016)

Ans. A cognitive method need to select which bandwidth and band frequency to use, and how long to send in this band. But, spectrum sensing only provides information about the spectrum occupancy at a given point in time. Hence, it is essential for the cognitive technique to maintain a record of the history of spectrum usage and to have detailed models for the traffic statistics. To obtain the optimal statistics, a secondary system observes the environment at all times and at all possible frequencies. Although such approach might be prevented by energy and hardware constraints.

A normal, analytically tractable model, that considers every channel is in one of two states, and has a fix transition probability from one state to another, is known as Markov model. The cognitive radio either moves to a different band or ceases transmission if the current band in which it was operating has to be released. The spectrum handover should be done in such a way that the performance of the link is least affected. If a primary customer reappears and wish to use the spectrum, the secondary system has to cease transmission in a hierarchical system. An important parameter of the system is the time period for which a secondary system transmits in a particular band sensing again. This time period depends on the previously determined statistics of the primary channel.

Q.16. *Explain the non cooperative games in brief.*

Ans. A game in which no player can make enforceable contracts outside of those specifically modeled in the game is known as noncooperative game. Therefore, any cooperation must be self enforcing. There is no outside party that can enforce a specific policy. An operating point where no customer has a motivation to unilaterally change its current strategy is known as Nash equilibrium. Nash equilibrium is an significant quantity in such type of games. An example of noncooperative games is a heavily loaded ALOHA system. There will be more number of collisions when every user transmits packets at a very high rate, and it can result in system congestion. But, a customer that endeavour to behave in a socially responsible way and sends out fewer packets will get an even lower throughput as long as the other customers keep up their high channel access rate.

Q.17. Explain in detail games with partial coordination.

Ans. Customers good behaviour may be enforced via different ways in cooperative games. The enforcement can be made via punishment by other players or through a central authority. Some of the most important methods are as follows –

(i) Referee Based Solution – With a referee enforcing specific kinds of behaviour the collisions and high interference conditions which plague standard competitive behaviour can be mitigated and the inefficiency of noncooperative games may be alleviated. The intervention level relies on the information level which the referee has and the control information that is to be transmitted to the customers. Fig. 4.5 shows a low overhead referee algorithm.

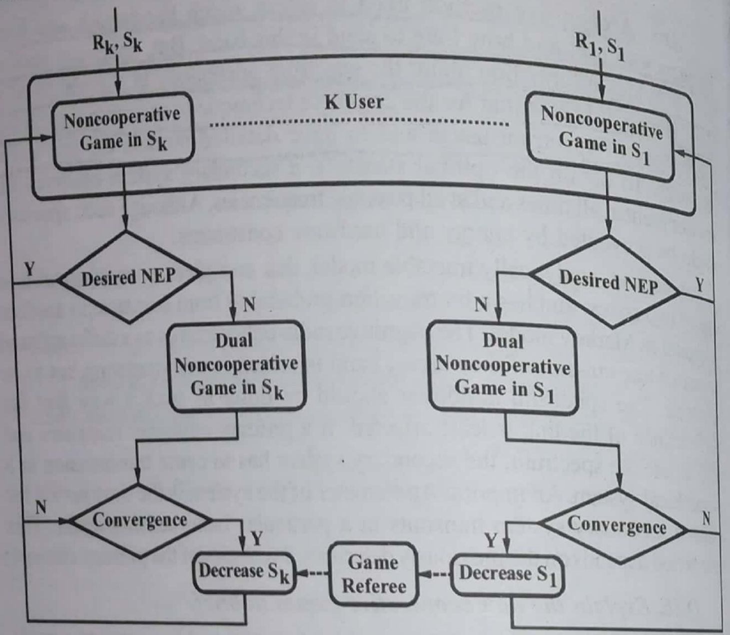


Fig. 4.5

(ii) Threat and Punishment in Repeated Games – Generally, the Nash equilibrium considers a static game, it means the players play only one time. However, in repeated games better results can be obtained, because players can learn from the past. As a result, a player will not sacrifice its long term benefits when it is going to be severely punished by the other players. The total payoff weighted average over time can be written as,

$$V = \sum_{t=1}^{T} \beta^{t-1} u_t \quad \text{...(i)}$$

Here, u_t = Payoff at time t
β = Discount factor
T = Total time.

(iii) Spectrum Auction – In a cognitive radio system, the spectrum auction is similar to any other auction in real life. The users are bidding a fixed price for the available spectrum, and a player who pays more will find a higher assignment of resources. The price need not be actual money and it may based on the quality of the user channel. For instance, it has been proposed that every player is charged according to SINR or received power.

(iv) Bargaining Solutions – Bargaining is another assignment algorithm which draws its inspiration from economic theory. Many players can negotiate with each other about who gets which spectral resources allocated, in this approach. Two neighboring players may change channels in one-to-one bargaining. An exchange of channels might be useful for both players because channel qualities are not essentially the same for various players. A bargaining with one buyer and multiple sellers may be applied when more players are involved.

Q.18. Write short note on centralized solutions.

Ans. In centralized solutions a central authority has all the information regarding the bandwidth needs of the several users as well as the available bandwidth and allocates the resources in such a manner to optimize the system spectral efficiency. It is considered that the secondary users follow commands from the central authority. A high overhead is involved in such a solution for transmitting all the needed information to the central authority and for the channel control information.

The spectrum allocation can be written as an optimization problem in following manner,

$$\min_{\mathbf{x}\in\Omega} f(\mathbf{x})$$
$$g_i(\mathbf{x}) \le 0 \qquad \text{for } i = 1, \ldots\ldots I$$
$$h_j(\mathbf{x}) = 0 \qquad \text{for } j = 1, \ldots\ldots J$$

where $\mathbf{x}$ = Parameter vector for the optimization of f() and has to lie in the space Ω

$g_i(\mathbf{x})$ = Inequality constraints for the parameter vector

$h_j(\mathbf{x})$ = Equality constraints for the parameter vector.

The meaning of these quantities can differ depending on the optimization goal and the application. For example **x** may be a vector having bandwidth assignment and power for the various users. Different users' power and bandwidth allocated are limited by the inequality constraints. For example, consider a condition with multiple secondary systems i, where i-th system gives bandwidth b_{ij} to the j-th user. Then a user sends its data to the i-th base station using that bandwidth. The data rate of a link assuming communication at Shannon capacity, is

$$R_{ij} = b_{ij} \log_2 \left[1 + \frac{|h_{ij}|^2 P_{ij}}{b_{ij}} \right]$$

Here, $|h_{ij}|^2$ = Channel gain

P_{ij} = Power on each link.

Hence, the optimization problem can be written as,

$$\max_{b_{ij}, P_{ij}, x_i} \sum_j \sum_i R_{ij}$$

$$\sum_j b_{ij} \le B_i$$

$$\sum_i P_{ij} \le P_j$$

$$\sum_i B_i \le B_{tot}$$

$$b_{ij} \ge 0,\ P_{ij} \ge 0,\ B_i \ge 0$$

These problems can be solved by standard Lagrangian optimization methods.

Q.19. *How is spectrum sharing done in cognitive radio ?*

(R.G.P.V., June 2014)

Or

How is spectrum sharing done in cognitive radio system ? Explain with the help of suitable block diagram. ***(R.G.P.V., June 2016)***

Ans. Game theory is an important mathematical tool in the analysis of spectrum sharing. A game made up of a number of players, a strategy for every player, and a payoff. Every user adjusts its strategy in such a way that the payoff is maximized. The optimum strategies depend on the coordination amount, and whether there is a central authority that can enforce certain rules of the games.

Sometimes, the ideal outcome of a game is to obtain a Pareto optimum (also called social optimum). In this case, there is by definition no other outcome that has a better payoff for at least one player, and at the same time does not minimize the payoff for any other player.

Also refer to the ans. of Q.16, Q.17 and Q.18.

Q.20. *Discuss about spectrum sensing, spectrum management and spectrum sharing.* ***(R.G.P.V., June 2017)***

Ans. **Spectrum Sensing** – Refer to the ans. of Q.11.

Spectrum Management – Refer to the ans. of Q.15.

Spectrum Sharing – Refer to the ans. of Q.19.

Q.21. *What do you mean by overlay ? Explain in detail.*

Ans. Overlay is another approach to cognitive radio. In overlay we do not try to prevent transmission at the same frequency and time by primary and secondary user rather, we use it. The secondary user can help the primary user by using suitable types of cooperation and still at the same time send its own information.

Consider, a primary and a secondary receiver/transmitter pair tries to communicate over an AWGN channel. Further consider that the secondary transmitter has enough message knowledge of the primary user.

The latter consideration is unrealistic in mostly all wireless setting, however it is reasonable in the two cases, which are as follows –

(i) A particular code is used by the primary user, that permits a receiver to decode the message as soon as it has gathered sufficient mutual information.

(ii) The primary user, uses automatic repeat request.

Now, a secondary transmitter can pursue one of the following two possible strategies –

(i) Selfish Approach – The secondary transmitter uses all its power to send the secondary message. Secondary transmitter uses its knowledge of primary message to make sure that there is no effective interference to the secondary receiver.

(ii) Selfless Approach – In this approach, only part of its power is used by the secondary transmitter to help convey the primary message to the primary receiver. The rest of the power is used to transmit the secondary message to the secondary receiver. The secondary system can really boost the rate of the primary message depending on the fraction of power used for transmitting the primary message. If the secondary system makes sure that the primary rate remains unchanged, then the primary system will be fully oblivious to the fact that a secondary system is available at all and its capacity in an AWGN channel is

$$R_1 = \log_2 \left[1 + \frac{P_1}{N_0} \right]$$

and still the secondary system can send some of its data at a rate,

$$R_2 = \log_2 \left[1 + (1 - \alpha') \frac{P_2}{N_0} \right]$$

Scanned by CamScanner

$$\text{where,}\quad \alpha' = \left[\frac{\sqrt{P_1}\left(\sqrt{1+|h_{21}|^2\frac{P_2}{N_0}\left(1+\frac{P_1}{N_0}\right)}-1\right)}{|h_{21}|\sqrt{P_2}\left(1+\frac{P_1}{N_0}\right)}\right]^2$$

Here, we considered that $|h_{11}| = |h_{22}| = 1$ and $|h_{21}| < 1$

This means secondary system can send its self information for free i.e., without reducing the primary system rate as long as the crosstalk channel from the secondary transmitter to the primary receiver is weaker than the primary channel.

Q.22. Explain frequency regulations and transmit power constraints of ultra wide bandwidth signals.

Ans. In ultra wide bandwidth system communications, where signals have large bandwidth the underlay principle is used. This type of high bandwidth has the possibility of very high spreading factors, i.e. the ratio of the signal bandwidth to the symbol cost is extremely high. A spreading factor of 10^5 to 10^6 is obtained for transmission bandwidths of 5×10^8 Hz and 5×10^9 Hz, respectively for a typical sensor network application with 5 k symbol/s throughput. Spreading over such a high bandwidth means that the power spectral density of the radiation can be made very low while still maintaining good signal to noise ratio at secondary receiver. A narrowband receiver can only see a small part of the whole secondary transmit power within its own system bandwidth as shown in fig. 4.6. It shows that the interference to primary system is small.

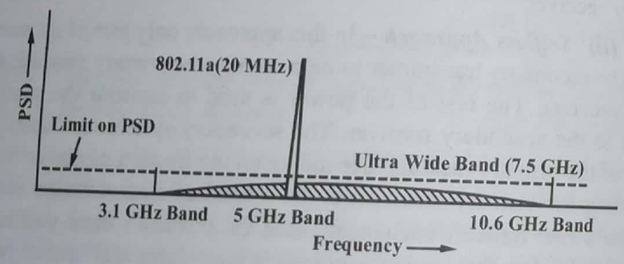


Fig. 4.6

Fig. 4.6 shows interference between a ultra wide bandwidth system and a narrowband local area network.

Frequency regulators defined ultra wide bandwidth signals as signals with a minimum of 500 MHz bandwidth. Frequency regulators stipulate that ultra wide bandwidth can be employed as unlicensed underlay systems as long as the power spectral density is limited to – 41.3 dBm/MHz. This power spectral density does not significantly disturb primary receiver, which is 10 m away from a UWB transmitter, because of its very low value. Ultra wide bandwidth's high spreading factor helps not only in mitigating interference to primary users but also enables a ultra wideband receiver to suppress narrowband interference by a factor which is equal to the spreading factor.

Q.23. What do you understand by overlay and underlay ?
(R.G.P.V., June 2014, 2016)

Ans. **Overlay** – Refer to the ans. of Q.21.

Underlay – The secondary users have severe restrictions on the transmit PSD in case of an underlay system. Then the effect on a primary receiver is a minor increase of the noise floor which the receiver sees. Such low PSD may be obtained by spreading the signal over a very large bandwidth or by keeping the transmit power low.

Also refer to the ans. of Q.22.

Q.24. How is spectrum sensing is achieved in cognitive radio ? Also define overlay and underlay. ***(R.G.P.V., June 2015)***

Ans. **Spectrum Sensing** – Refer to the ans. of Q.11.

Overlay and Underlay – Refer to the ans. of Q.21 and Q.23.

Q.25. Explain the methods of ultra wideband (UWB) signal generation in brief.

Ans. Various methods of ultra wideband signal generation are given below –

(i) Orthogonal Frequency Division Multiplexing – Refer to the ans. of Q.2, Unit-II.

(ii) Frequency Hopping – It uses various carrier frequencies at different times. The range of the oscillator determines the bandwidth of the resulting signal, not the bandwidth of the original signal which is to be transmitted. Implementation of a frequency hopping transmitter is simple. It is a traditional narrowband modulator followed by a mixer with the output of a frequency agile oscillator. There are two types of frequency hopping and there is a considerable difference in performance for both the systems – slow frequency and fast frequency hopping.

In SFH one or more data bits are transmitted within one hop, i.e. hopping rate is less than the message bit rate. In FFH one data bit is divided over multiple hop, i.e. frequency hopping rate is greater than message bit rate.

(iii) Direct Sequence Spread Spectrum – In a manner similar to FHSS, direct sequence spread spectrum can be thought of as a two stage modulation technique.

In the first stage, each transmitted information bit is spread into N smaller pulses referred to as chips. In the second stage, the chips are transmitted over a traditional digital modulator. At the receiver the transmitted chips are first demodulated and then passed through a correlator that despreads the signal. The DSSS can also be employed for code division multiple access. In a multiuser DS-CDMA environment, different codes are assigned to different users. In other words, each user has its own unique key code which is used to spread and despread only that user's messages.

(iv) Time Hopping Impulse Radio – Refer to the ans. of Q.33, Unit-I.

Q.26. Write the advantages of ultra wide bandwidth transmission.

Ans. The advantages of ultra wide bandwidth transmission are given below –

(i) The low power spectrum density and large spreading factor also give increased protection against eavesdropping.

(ii) A ultra wide bandwidth receiver may suppress narrowband interference by a factor which is nearly equal to the spreading factor.

(iii) A large absolute bandwidth can result in a high resilience to fading.

(iv) A large absolute bandwidth also leads to high accuracy of ranging and geolocation.

OO

UNIT 5

COOPERATIVE COMMUNICATION

INTRODUCTION AND MOTIVATION, FUNDAMENTALS OF RELAYING, RELAYING WITH MULTIPLE PARALLEL RELAYS

Q.1. Explain in brief about the principle of relaying.

Ans. Conventional wireless communications depend on point-to-point communication, that is, the communication of data is performed by only two nodes. These two nodes vary according to the communication medium, e.g., the base station and mobile station are employed in a cellular setting, laptop and access point are employed in wireless local area networks, and two mobile stations are employed in peer-to-peer communications. Other wireless receivers and transmitters which are in the proximity compete for the same resources, giving rise to interference. Information from the message source to the intended destination is received with the help of the following –

(i) Peer Nodes serving as Relays – This type of peer nodes (sensor nodes or mobile handsets) may change their roles based on the condition at hand. Sometimes they help to forward information and sometimes they works as a source.

(ii) Dedicated Relays – This type of relays can never work as source of the information, however they can facilitate the information exchange of other nodes.

In the system design relay nodes generate more degrees of freedom which can help to improve the performance, but also complicates the process of design. The cooperative communications is actually a logical network structure which results from the urge to get better coverage and efficiency. At beginning, assume a 3-node network in which only free space attenuation arises on every link as shown in fig. 5.1.

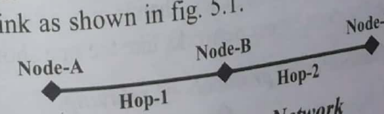


Fig. 5.1 Two-hop Network

Consider a condition in which node A does not contain sufficient power to transmit a data packet directly to node C. But, it can be done in two steps. In a first step, node A can send the data packet to the node B, which resends the packet and node C receives this resent packet. Extending this two hop approach to larger networks, the range which the network covers can be raised through multi-hopping. Fig. 5.2 shows a multi-hop ad-hoc network.

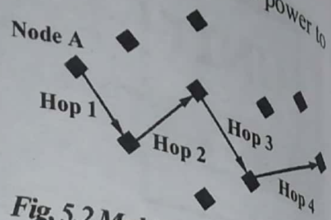


Fig. 5.2 Multi-hop ad-hoc Network

Broadcast effect is an important feature of the wireless propagation channel. If one node sends a signal, it can be collected by any node in the vicinity, this fact can be positively exploited in multinode networks. Multi-hopping does not use broadcast effect. It is used by more advanced cooperative communications approach. As shown in fig. 5.3(a), when node A sends, the signal reaches not only node B but also node C in weaker form. This weak signal might not be sufficient by itself for node C to decode, although it may be used to increase the signal obtained in a subsequent transmission from node B to node C.

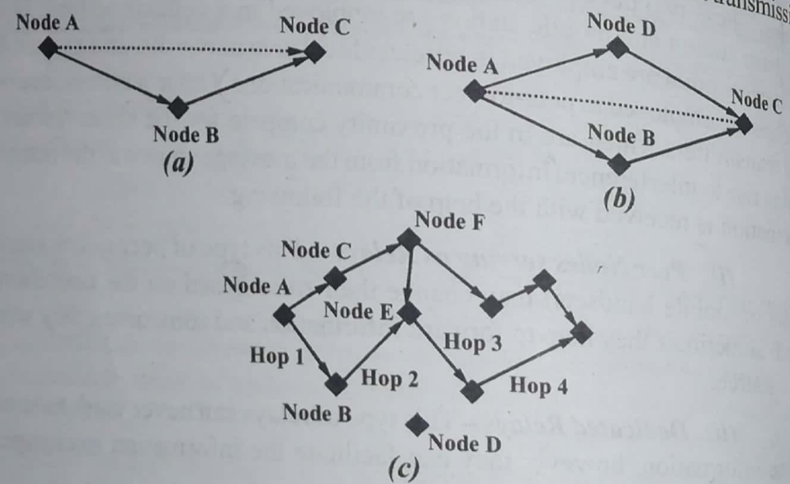


Fig. 5.3 Demonstrating the Broadcast Effect in Relay Networks

The broadcast effect has even more prominent effect in bigger networks. Consider the network shown in fig. 5.3 (b) if the first node sends, the signal arrives nodes B and D at same power. Reaching these two nodes does not need more transmit power. Now nodes B and D cooperate to forward the information to node C. This type of transmission is more efficient than if only a single node sends. The same principle is true for even bigger networks like the one shown in fig. 5.3(c).

Q.2. Explain fundamental protocols of relaying.

Or

What are the fundamentals of relaying in co-operative communication ?

(R.G.P.V., June 2015)

Ans. In a 3-node network the relay and destination are connected to a source by channels h_{sd}, h_{sr}, and h_{rd} between the nodes as shown in fig. 5.4.

By a fixed factor the relay amplifies the received signal and resends it in amplify-and-forward. The relay decodes the packet and then subsequently re-encodes and resends it in decode-and-forward. In compress-and-forward, the relay generates a quantized version of the signal it finds from the source and forwards that to the destination. The destination adds this compressed signal with the directly sent signal from the source. We consider that all relaying nodes cannot transmit and receive at the same time in the same frequency band. The transmit and receive levels of wireless signals are so different that the sent signal would swamp the receiver and make it impossible to found the received signal.

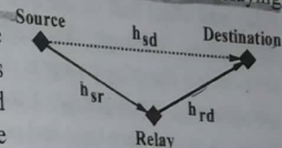


Fig. 5.4 Basic Relay Channel

Now these methods of relay processing can be added with different transmission protocols which prescribe when what information blocks are sent from which nodes. According to increasing performance, these methods are given below –

(i) Multi-hop xF³ – In this method only relay listenes and source sends, in first time slot. In the second time slot, destination is listening and only relay is sending.

(ii) Split Combine xF – In this method the source sends and only the relay listenes in the first time slot, but in second time slot, both source and relay send and the destination is listening.

(iii) Diversity xF – The source sends and both relay and destination are listening in first time slot. Only relay is sending and the destination is listening in second time slot.

(iv) Nonorthogonal Diversity xF – It is represented by NDxF. In this method the relay transmits differently encoded information in the second time slot.

(v) Intersymbol Interference xF – It is denoted by IxF. This method works only when the relay has full duplex capability. The source transmits an information to the relay in the ith time slot, and the relay forwards this information to the destination. On the other hand, the source transmits the next information to the relay at the same time. The destination is continuously listening and in every times lot listens the superposition of the current information block directly from the source and a previous information block from the relay.

Q.3. Discuss three node relay channel model. **(R.G.P.V., Dec. 2016)**

Ans. Refer to the ans. of Q.2.

Q.4. Give the classification of fundamental protocols.

Ans. The classification of fundamental protocols can be done according to the following criteria –

(i) Fixed versus Adaptive Allocation of Time and Spectral Resources – The most usual form of half duplex protocols considers that the network spends half the time transmitting a packet from the source to the relay and other half of the time forwarding that packet from the relay to the destination. Although, this might not be the most effective use of the given time. When the relay power is constant, then transmission should be done at a rate close to the capacity of the relay destination channel. When the relay power is varied then time and power should be optimized simultaneously.

(ii) Fixed versus Adaptive Power – The power can be expended by the source and by the relay fixed or it can be adaptive to the channel state. Power is adapted, if it is adaptive, over the nodes, over time or over frequency. Constraints are then generally imposed on the sum transmit power in which the summation is over nodes, time and frequency.

Q.5. What is degraded relay channel ? **(R.G.P.V., Dec. 2016)**

Ans. The relay channel is degraded if

$$p(y_r, y_d \mid x_s, x_r) = p(y_r \mid x_s, x_r)\, p(y_d \mid y_r, x_r) \quad ...(i)$$

The degraded relay channel defined in equation (i) can be interpreted as the case where the destination receives a combination of the degraded version of the signal observed at the relay Y_r and the output of the relay X_r. More specifically, the relay channel is degraded if $p(y_d \mid x_s, x_r, y_r) = p(y_d \mid x_r, y_r)$, i.e. $X_s \rightarrow (X_r, Y_r) \rightarrow Y_d$ is a Markov chain.

Q.6. Explain in detail decode and forward.

Ans. Decode and forward is a very important relaying scheme. A packet is received by the relay which decodes it, hence mitigating the effects of noise before resending and re-encoding the packet. Multi-hop decode-and-forward is a very simple method to analyze the capacity of several of the implementations. The overall data rate per unit bandwidth can be defined, if we consider a given transmit power and equal dividing of the available time between the two phases as –

$$R = \frac{1}{2}\text{Min}\left[\log\left(1+\frac{P_s|h_{sr}|^2}{P_n}\right), \log\left(1+\frac{P_r|h_{rd}|^2}{P_n}\right)\right] \quad ...(i)$$

where,
P_n = Noise power
P_s = Source power
P_r = Relay power.

The terms inside the min operation are the capacities of the relay-destination and source-relay link. The values of P_r and P_s can be constant or can be optimized given the power constraints and the channel coefficients h_{rd} and h_{sr} values. Hence

$$P_r = P_0\frac{|h_{sr}|^2}{|h_{sr}|^2+|h_{rd}|^2}$$

$$P_s = P_0\frac{|h_{rd}|^2}{|h_{sr}|^2+|h_{rd}|^2} \quad ...(ii)$$

The destination hears during both the phases and hence can sum the signals it obtained from the source and the relay in diversity decode and forward scheme. Then we define two important cases as follows –

(i) Transmission from the Relay using Repetition Coding – Same encoder is used by both relay and source in this case. For improving the SNR the destination can sum up the received signals before the decoding. Now consider the further limitation, transmission is possible only when the destination perfectly receives the message from both source and relay. Hence, the optimum achievable rate will be as follows –

$$R_{DDF} = \frac{1}{2}\text{Min}\left[\log\left(1+\frac{P_s|h_{sr}|^2}{P_n}\right), \log\left(1+\frac{P_r|h_{rd}|^2}{P_n}+\frac{P_s|h_{sd}|^2}{P_n}\right)\right] \quad ...(iii)$$

and its maximum power allocation is

$$P_s = P_0\frac{|h_{rd}|^2}{|h_{sr}|^2+|h_{rd}|^2-|h_{sd}|^2} \quad ...(iv)$$

$$P_r = P_0\frac{|h_{sr}|^2-|h_{sd}|^2}{|h_{sr}|^2+|h_{rd}|^2-|h_{sd}|^2}$$

(ii) Transmission from the Relay using Incremental Redundancy Encoding – The relay decodes the packet and re-encodes it with a different coder in this case. The receiver can add up the mutual information from the two transmission phases intuitively. Hence the capacity of such a protocol can be defined as follows –

$$R_{DDF,IR} = \frac{1}{2}\text{Min}\left[\log\left(1+\frac{P_s|h_{sr}|^2}{P_n}\right), \log\left(1+\frac{P_r|h_{rd}|^2}{P_n}\right)+\log\left(1+\frac{P_s|h_{sd}|^2}{P_n}\right)\right] \quad ...(v)$$

We can explain the entire system to be in outage using the concept of outage capacity if the achievable rate falls below the desired threshold R_{th}. The probability for such an outage for nonadaptive DDF can be defined as –

$$\Pr(R_{DDF} < R_{th}) = \Pr\left[|h_{sr}|^2 < (2^{2R_{th}} - 1)\frac{P_n}{P_s}\right] + \Pr\left[|h_{sr}|^2 > (2^{2R_{th}} - 1)\frac{P_n}{P_s}\right] \Pr[|h_{sd}|^2 P_s + |h_{rd}|^2 P_r < (2^{2R_{th}} - 1) P_n] \quad ...(vi)$$

Here the first term on the right hand side corresponds to the case in which source-relay connection is weak and second term corresponds to the case in which source-relay connection is strong but source-destination and relay-destination connections are weak to sustain enough information flow to the destination.

Q.7. Explain the following in brief –

(i) Amplify-and-forward (ii) Compress-and-forward.

Ans.* (i) *Amplify-and-Forward – The relay takes the noisy signal y_r that it receives and amplifies it with a gain β, is the basic principle of amplify and forward.

Consider that the amplify and forward processing at the relay causes a delay of half a time slot. Hence, the signal achieved at the destination is purely the signal from the source, plus noise, in first phase. The signal is the sum of the direct signal from the source, and the signal from the relay, that is the amplified source word of the previous phase of that slot, in second phase. Hence,

$$y_d^{(2)} = h_{sd}x_s^{(2)} + h_{rd}x_r^{(2)} + n_d^{(2)}$$
$$= h_{sd}x_s^{(2)} + \beta h_{sr}h_{rd}x_s^{(1)} + \beta h_{rd}n_r^{(1)} + n_d^{(2)} \quad ...(i)$$

where x_r = Relay transmit signals
x_s = Source transmit signals
n_d = Destination noise
n_r = Relay noise.

Superscripts $^{(1)}$ and $^{(2)}$ represent the first and second transmission phase respectively.

Power constraints limit the amplification factor. In instantaneous power constraint case, we need that

$$|\beta|^2 \le \frac{P_r}{P_n + P_s|h_{sr}|^2} \quad ...(ii)$$

We first determine the performance of multi-hop amplify-and-forward with $P_r = P_s = P$. At the receiver, the signal reaching during the first phase is simply the source signal multiplied with the channel coefficient h_{sd} and the noise has variance P_n.

The signal reaching at the destination is the source signal multiplied with $\beta h_{sr}h_{rd}$ in the second phase and the noise has variance,

$$P_n' = (|\beta|^2 |h_{rd}|^2 + 1) P_n \quad ...(iii)$$

The maximum power allocation in MAF can be defined as –

$$\frac{P_s}{P_r} = \sqrt{\frac{|h_{rd}|^2 P_0 + P_n}{|h_{sr}|^2 P_0 + P_n}} \quad ...(iv)$$

To maximize the SNR, in diversity-amplify-and-forward (DAF), signals from the two phases need to be combined. Hence the signal from the first phase has to be multiplied by

$$\sqrt{\frac{P_s}{P_n}} h_{sd}^* \quad ...(v)$$

The signal from the second phase has to be multiplied by

$$\sqrt{\frac{P_s\beta^2}{P_n'}}(h_{sr}h_{rd})^* \quad ...(vi)$$

Considering $P_s = P_r = P$, the corresponding SNRs for the two phases are as follows –

$$\gamma_1 = \frac{P|h_{sd}|^2}{P_n} \quad ...(vii)$$

$$\gamma_2 = \frac{\left(\frac{P}{P_n}\right)^2 |h_{sr}h_{rd}|^2}{\frac{P}{P_n}|h_{sr}|^2 + \frac{P}{P_n}|h_{rd}|^2 + 1} \quad ...(viii)$$

(ii) Compress-and-Forward – Compress-and-forward is similar to amplify-and-forward in the manner that the relay does not decode a message, but forwards whatever it receives. However, the forwarded message is a quantized, compressed version of the message received at the relay, in CF. Compress-and-forward has been shown to give greater capacity than DF or AF for some particular configurations of channel.

Q.8. Write short note on the relaying with multiple parallel relays.
(R.G.P.V., June 2014)

Ans. When more than one relay are available for forwarding the information, a cooperation between several relays may be used to highly improve the performance of the relaying method specially in fading channels. The multiple relays give diversity ways which bring robustness with respect to fading and interference. The cooperation between the relays requires the exchange of channel state information and control information. Hence there are many methods which trade off the overhead with the performance of system in various ways. Two phase transmission with parallel relays is shown in fig. 5.5. Transmission takes place in two phases – in first phase the source broadcasts the information and in second phase, one or more relays forward the information to the destination. Phase one takes the merit of the broadcast effect. The second phase particularly the transmission from a multiantenna transmitter to a single antenna destination resembles a smart antenna system. The main difference between the two is that the antennas are distributed over a large area in space.

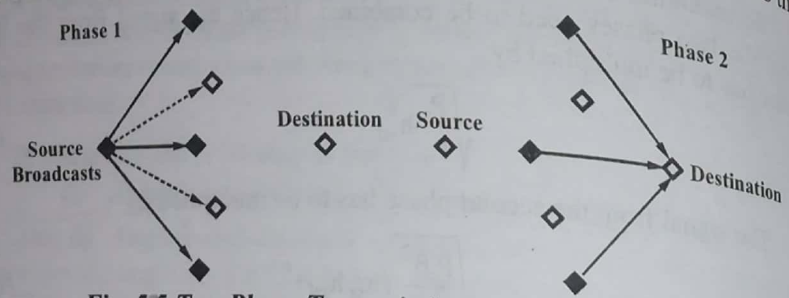


Fig. 5.5 Two Phase Transmission with Parallel Relays

In first phase, always consider that the k-th relay knows the channel from the source to it. If the source can either change its power or its transmission time according to the channel states only then channel state information at the TX is useful. Again consider CSIR for second phase. Different transmission methods used depending on the type of channel state information at the transmitter are given below –

(i) Full CSIT Available – These relays know both phase and amplitude of the channel to the destination. In this case virtual beamforming can be used. This method ensures the maximum SNR available at the receiver for a given sum power expenditure at the relays.

(ii) Amplitude CSIT Available – In this case relay do not know the phase but know the amplitude of the channel to the destination. The best strategy is to choose a single relay which gives the better quality of transmission.

(iii) No CSIT Available – The relays can send data packet encoded in space time. They work like antennas in a transmit diversity system without CSIT.

(iv) Average CSIT Available – Only the mean channel gain is available in this case not the instantaneous realization.

Q.9. Describe multi-hop wireless communication. (R.G.P.V., May 2018)

Ans. Refer to the ans. of Q.8.

Q.10. Explain following transmission schemes –
(i) Relay selection (ii) Distributed beamforming.

Ans. (i) Relay Selection – Selecting the best of all the available relays is referred as relay selection. This selection has following challenges –

(a) Defining what do you mean by best relay

(b) Really finding the relay for a given set of channel states.

For defining best relay, consider two cases – if the source has certain transmit power and data rate then we cannot influence which relay will get the packet correctly during phase first transmission otherwise we can select the set of relays which receive the packet and choose the one for forwarding which has the strongest channel to the destination. We can make sure that a particular relay gets the message in first phase when source can adapt to the channels. Selecting this relay needs a balancing between the source relay and relay-destination channel strengths. We can choose the relay which gives the best,

$$\eta_k = \text{Min } [|h_{s,k}|^2, |h_{k,d}|^2] \qquad ...(i)$$

Another criterion assumes a smoothed out version of this criterion

$$\eta_k = \frac{2}{\dfrac{1}{|h_{s,k}|^2} + \dfrac{1}{|h_{k,d}|^2}} \qquad ...(ii)$$

Frequently, a central control node knows all the $|h_{s,k}|^2$ and $|h_{k,d}|^2$. Algorithms which can be used for this purpose are –

(a) The destination transmits out a brief broadcast signal which permits the relays to calculate the $|h_{k,d}|^2$ in first step.

(b) The source transmits the data packet as well as a CTS message after it is finished. Every relay tries to get the packet and also calculates its $|h_{s,k}|^2$.

(c) Every relay begins a timer with an initial value K_{timer}/η_k and counts down when listening for possible on-air signals from the other relays. The relay begins to send when the timer reaches 0, unless another relay has already began to send. For K relays, the outage probability is determined for the case of Rayleigh fading on all links as follows –

$$\Pr[I < R_{th}] = \prod_{k=1}^{K} \left[1 - \exp\left[-\frac{2^{2R_{th}} - 1}{P/P_n} \left(\frac{1}{\bar{\gamma}_{s,k}^2} + \frac{1}{\bar{\gamma}_{k,d}^2} \right) \right] \right] \qquad ...(iii)$$

where $\bar{\gamma}_{k,d}^2$ = Mean channel gains of the relay-destination channel

$\bar{\gamma}_{s,k}^2$ = Mean channel gains of the source-relay channel.

(ii) Distributed Beamforming – The source broadcasts the informations and a set D gets the packet in good order in first phase. The optimal transmission coefficient at every chosen relay k for an MDF protocol can then shown to be proportional to

$$\frac{h^*_{k,d}}{(\Sigma_{k \in D}|h_{k,d}|^2)^{1/2}} \quad \text{...(iv)}$$

once the channel state information at the transmitter is known.

The optimum gain applied at relay k for relays using amplitude and forward is follows –

$$w_k = K^{AF} \frac{|h_{sk}||h_{kd}|}{1+P_s|h_{sk}|^2+P_k|h_{kd}|^2} \frac{h^*_{sk}}{|h_{sk}|} \frac{h^*_{kd}}{|h_{kd}|} \quad \text{...(v)}$$

The power on the k-th relay is as follows –

$$P_k \propto \frac{|h_{sk}|^2|h_{kd}|^2[P_s|h_{sk}|^2+1]}{[1+P_s|h_{sk}|^2+P_k|h_{kd}|^2]^2}$$

The CSIT determination at the relays is nontrivial i.e., not only does every relay require to know its channel to the destination but it also requires to know the sum of the channel gains from all the relays which will be active in the data forwarding.

Q.11. Discuss the transmission process on orthogonal channel in brief.

Ans. When channel state information is unavailable at the transmitter, then each relay can send an orthogonal channel. Obviously, this technique removes the interference between the different relay channels. Although, it also causes a drastic reduction of the spectral efficiency. Particularly, assume a diversity-decode-and-forward (DDF) technique in which every relay has a reserved channel, whether it can decode the message or not. Its capacity can be defined as follows –

$$I = \frac{1}{K+1}\log\left[1+\gamma_{s,d}+\sum_{k \in D}\gamma_{k,d}\right] \quad \text{...(i)}$$

where D denotes the relays set which can decode the message from a certain source. When all links are Rayleigh fading, the outage probability conditioned on a certain decoding set for this technique is for large signal-to-noise ratio,

$$\Pr[I < R_{th}|D] \sim [2^{(K+1)R_{th}}-1]^{|D(s)|+1} \frac{1}{\gamma_{s,d}} \prod_{k \in D}\frac{1}{\gamma_{k,d}} \frac{1}{[|D|+1]!} \quad \text{...(ii)}$$

The probability of determining a particular decoding set can be defined as –

$$\Pr[D] \sim [2^{(K+1)R_{th}}-1]^{K-|D(s)|} \prod_{k \notin D}\frac{1}{\gamma_{s,k}} \quad \text{...(iii)}$$

Then, the overall outage probability is equation (ii) unconditioned by equation (iii). This expression is defined as follows –

$$\left[\frac{2^{(K+1)R_{th}}-1}{\bar{\gamma}^{lb}}\right]^{K+1} \sum_{k \in D}\frac{1}{[|D|+1]!} \lesssim \Pr[I < R_{th}] \lesssim$$

$$\left[\frac{2^{(K+1)R_{th}}-1}{\bar{\gamma}^{ub}}\right]^{K+1} \sum_{k \in D}\frac{1}{[|D|+1]!} \quad \text{...(iv)}$$

where $\frac{1}{\bar{\gamma}^{lb}_k} = \text{Min}\left\{\frac{1}{\bar{\gamma}_{s,k}}, \frac{1}{\bar{\gamma}_{k,d}}\right\} \frac{1}{\bar{\gamma}^{ub}_k} = \text{Max}\left\{\frac{1}{\bar{\gamma}_{s,k}}, \frac{1}{\bar{\gamma}_{k,d}}\right\} \bar{\gamma}^{lb}_s$

$$= \bar{\gamma}^{ub}_s = \bar{\gamma}_{s,d}$$

and $\bar{\gamma}^{lb}$ denotes the geometric mean of the $\bar{\gamma}^{lb}_k$, k = 1... K+1, and similarly for $\bar{\gamma}^{ub}$.

Let us consider the situation in which there are multiple nodes and all of them serve as sources, or relays and that can send at different frequencies or times. Fig. 5.6 shows the situation in which each node is sending information from a particular source only for 1/(K+1) of the available time.

Phase I	Phase II			
1 Sends	2 Repeats 1	3 Repeats 1	. . .	K + 1 Repeats 1
2 Sends	1 Repeats 2	3 Repeats 2	. . .	K + 1 Repeats 2
3 Sends	1 Repeats 3	2 Repeats 3	. . .	K + 1 Repeats 3
⋮	⋮	⋮	⋮	⋮
K+1 Sends	1 Repeats K + 1	2 Repeats K + 1	. . .	K Repeats K + 1

Frequency

Time

Fig. 5.6

Q.12. Write short note on distributed space-time coding.

Ans. We assume the situation in which the information is transmitted from source to the relays in first phase. Now, a space-time coded transmission to the destination is performed by the relays in second phase. Alternatively, each relay node serves as a virtual antenna and transmits the signal that, in a MIMO setting, would be transmitted by one of transmit antenna array elements. For example, the used space-time code could be an Alamouti code, if two relays are used. It implies that two symbols (c_1 and c_2) are sent from the two relays at first time instant. Then,

$$s_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \quad ...(i)$$

where s denotes the vector comprising the symbols transmitted from the relays. The signal vector at second time instant

$$s_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -c_2^* \\ c_1^* \end{pmatrix} \quad ...(ii)$$

is sent. The communication protocol must have a way to allocate to each relay which antenna it is, and consequently, which data sequence ($c_1 - c_2^*....$) or ($c_2c_1^*....$) it should transmit.

As the Alamouti code is a rate-1 code, the transmission spectral efficiency is better than for the relaying on orthogonal channels, in which the rate is only half. When using more relays, the relaying spectral efficiency with orthogonal space-time codes reduces somewhat i.e., for K > 2, no rate-1 orthogonal space-time codes exist. For K = 3 or 4, the achievable rate reduces to 3/4. However, this is better than orthogonal relaying, in which the rate reduces as 1/K.

Q.13. What is space time code ? ***(R.G.P.V., Dec. 2016)***

Ans. Refer to the ans. of Q.12.

Q.14. Are there any other advantages than capacity for co-operative communication ? If yes, list. ***(R.G.P.V., Dec. 2016)***

Or

What are the benefits of co-operative transmission ? (R.G.P.V., May 2018)

Ans. The main advantages of using supportive, cooperative or space-time relays in the system are given below –

(i) Performance Gains – Large system-wide performance gains are obtained due to pathloss gains as well as diversity and multiplexing gains. These translate into decreased transmission powers, higher capacity or better cell coverage.

(ii) Reduced Costs – Compared to a purely cellular approach to provide a given level of QoS to all users in the cell, relaying is a more cost effective solution. Although, it has also been represented that whilst savings are not as dramatic as hoped for, the capital and operational expenditures are usually lower when relay are used.

(iii) Balanced Quality of Service – Whilst in traditional systems users at the cell edge or in shadowed areas suffered from capacity or coverage problems, relaying permits to balance this discrepancy and hence give (almost), equal quality of service (QoS) to all users.

(iv) Infrastructure-less Deployment – The relays are used to permits the roll-out of a system which has minimal or no infrastructure available prior to deployment. For instance, in disaster-struck areas, relaying can be used to facilitate communications even though the cellular system is nonfunctioning. For hybrid deployments, that is a cellular system coupled with relays, the deployment and maintenance costs can be lowered.

Q.15. Explain the principle of coded cooperation in brief.

Ans. In this cooperation, relaying and error correction coding are integrated, resulting in improved diversity. The principle of coded cooperation is shown in fig. 5.7. At node 1, a source data block is encoded with a forward error correction, and the resulting codeword is divided into two parts, with N_1 and N_2 bits respectively. We can reconstruct the source data from the first N_1 bits alone.

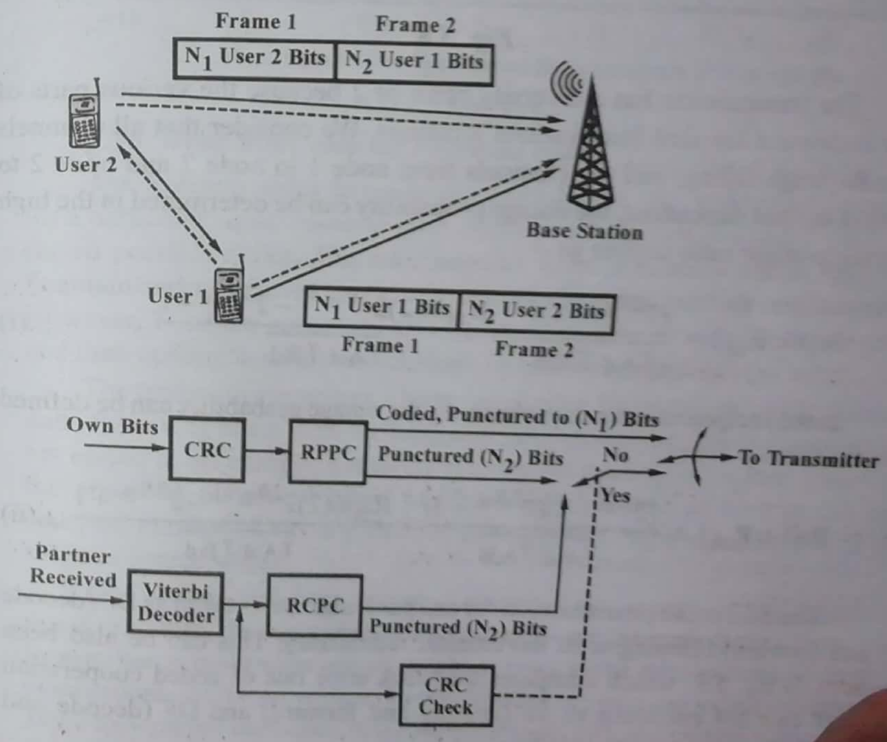


Fig. 5.7

Now the transmission interval of time available for one block of source date is splitted into two parts, during the first interval of time, node 1 broadcasts the first N_1 bits. They are received by the destination and also by node 2. At the same time, node 2 sends its first N_1 bits. If node 1 can completely decode the source word of node 2, then it determines the second N_2 bits assigned with that source data of node 2 and sends it in the second interval of time. If node 1 cannot decode completely, then it transmits the N_2 bits along with its own codeword. Node 2 also acts in a similar manner. As there is no feedback between nodes 1 and 2, the four conditions can originate as shown in fig. 5.8.

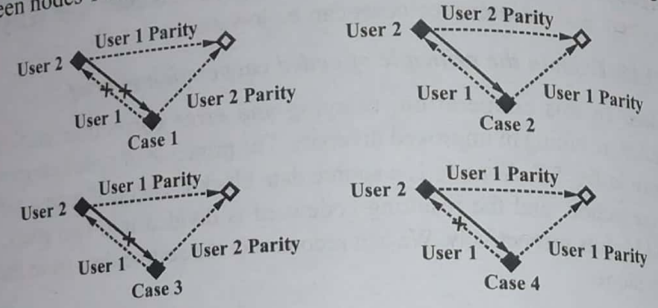


Fig. 5.8

The transmission has a diversity order of 2 because the various parts of the codeword are sent from various locations. We consider that all channels are Rayleigh fading, and the channels from node 1 to node 2 and node 2 to node 1 are not dependent, the outage probability can be determined in the high signal-to-noise ratio regime as –

$$\Pr[I < R_{th}] = \frac{(2^{2R_{th}}-1)^2}{\gamma_{A,d}\,\gamma_{A,B}} + \frac{R_{th}\,ln(2)2^{2R_{th}+1} - 2^{2R_{th}} + 1}{\gamma_{A,d}\,\gamma_{B,d}} \quad \text{...(i)}$$

In the reciprocal inter-user channels, the outage probability can be defined as follows –

$$\Pr[I < R_{th}] = \frac{(2^{R_{th}}-1)(2^{2R_{th}}-1)}{\gamma_{A,d}\,\gamma_{A,B}} + \frac{R_{th}\,ln(2)2^{2R_{th}+1} - 2^{2R_{th}} + 1}{\gamma_{A,d}\,\gamma_{B,d}} \quad \text{...(ii)}$$

The data packet transmission for one particular user is taken as DF (decode and forward) relaying with incremental redundancy. This can be also been seen in fig. 5.9, which compares the block error rate of coded cooperation with rate 1/4 encoding to AF (amplify and forward) and DF (decode and forward) with rate 1/2 encoding.

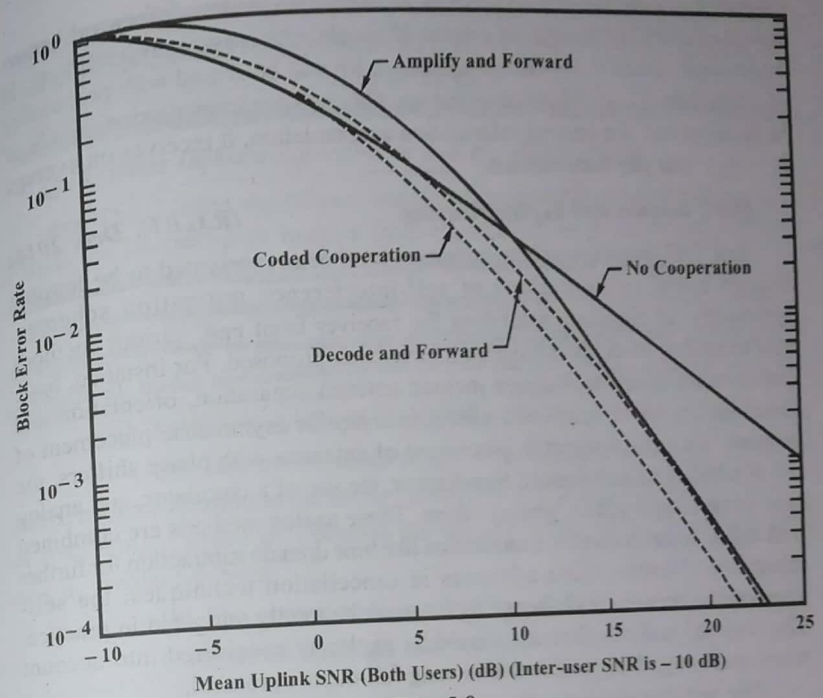


Fig. 5.9

Q.16. Write short note on Fountain codes.

Ans. Rateless codes do not have a fixed coding rate, and are not optimized for a particular signal-to-noise ratio. More precisely, rateless codes work well for all possible SNRs. The most popular form of rateless codes, known as Fountain codes, are used to prevent the erasure of data packets on the Internet. However, Fountain codes can also be made to operate on a bit-by-bit basis, and then operate in an erasure channel in following manner –

The transmitter generates a bitstream from a finite-length block of source data. Receiving nodes view the bitstream, and accumulate the bits which were not erased in the channel. Receivers can recover the original information from the observed, unordered subset of the codestream, just as long as the total received number of bits is greater as compared to the number of bits in the original source word.

As Fountain codes are universal codes which operate at any signal-to-noise ratio rate, the same code design is employed for transmitting from one transmitter to multiple receivers in which links to the transmitter have many attenuations. At the same time, Fountain codes can be used in relay networks since they permit each node in the network to accumulate mutual information from multiple sending nodes. Intuitively, the difference between mutual-information accumulation and energy accumulation is straightforward known from the binary signaling over erasure channels with erasure probability p_e. If the receiver accumulates energy, then each bit will be erased with probability p_e^2, therefore $1-p_e^2$ bits on average are obtained per transmission. While, if the receiver can use mutual-information accumulation, it receives on average $2(1-p_e)$ bits per transmission.

Q.17. Discuss full duplex relaying. ***(R.G.P.V., Dec. 2016)***

Ans. Full-duplex wireless operation has been represented to be feasible through novel combinations of self-interference mitigation schemes. Particularly, to eliminate saturating the receiver front end, various methods prior to analog-to-digital conversion have been proposed. For instance, basic analog cancellation techniques include antenna separation, orientation and directionality. More involved techniques comprise asymmetric placement of transmit antennas, symmetric placement of antennas with phase shifters, the use of a balanced/unbalanced transformer, the use of a circulator, and analog time domain subtraction, among others. These analog methods are combined with digital techniques after quantization like time domain subtraction for further mitigation. Despite these advances in cancellation techniques, the self-interference remains a challenge as it cannot be totally mitigated in practice. The residual self-interference should be explicitly considered into account when assessing, designing and analyzing full-duplex protocol.

The full-duplex protocols are divided depending on whether the direct source – destination link is used for transmission. In fact, the idea of cooperative relaying can be traced back to the works of Vander Meulen and Cover in which make use of the direct link for full-duplex communication. Particularly, in cooperative full-duplex schemes, the source transmits to the relay and the destination, while the relay simultaneously receives the signal from the source and transmits to the destination, as shown in fig. 5.10. The direct link is not used in FDDH protocols. FDDH relaying has two important disadvantages when the direct link is not under heavy shadowing. First, although the source is permitted to transmit continuously, the rate of the FDDH scheme might be degraded due to the self-interference created at the destination node from the direct link. Moreover, this protocol does not provide any diversity benefits. Hence, cooperative full-duplex methods that make use of the direct link for transmission might be able to give rate and diversity benefits.

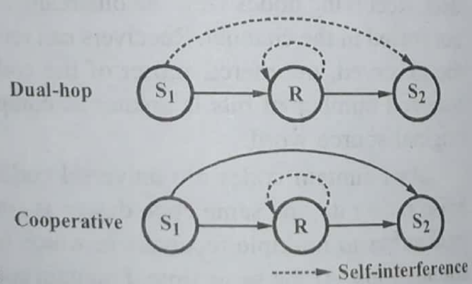


Fig. 5.10 Full-duplex Relay Protocols

ROUTING AND RESOURCE ALLOCATION IN MULTI HOP NETWORKS

Q.18. Write Dijkstra algorithm to find the shortest path.

Ans. The Dijkstra algorithm is based on the principle of greedy relaxation. It finds the shortest-path weights from a source node to every node in the network. The steps involved in this algorithm are as follows –

(i) Assign the source node the distance $d_s = 0$, and all other nodes the distance $d_i = \infty$. In addition, declare the source node the "current node" and mark all nodes as "unvisited".

(ii) Loop until all nodes are visited –

(a) Take all unvisited neighbor nodes i of the current node c, that is, nodes with links to the current node. For each neighbor node i, find the $\hat{d}_i = d_c + w_{c,i}$ where $w_{c,i}$ denotes the edge cost between nodes c and i. If $\hat{d}_i < d_i$, replace d_i by $\hat{d}_i$.

(b) The current node is marked as visited.

(c) Take the unvisited node with the smallest distance as current node.

This is an efficient algorithm. Based on the particular implementation the runtime is proportional to $O(|V|^2 + |E|)$ or $O(|V| \log|V| + |E|)$, where V denotes the number of vertices, and |E| denotes the number of edges.

Q.19. Write Bellman-Ford algorithm to find shortest path.

Ans. Bellman-Ford algorithm also finds the shortest path. The steps involved in this algorithm are as follows –

(i) Assign the source node the distance $d_s = 0$, and all other nodes the distance $d_i = \infty$.

(ii) Repeat |V| – 1 times.

(a) For all edges in the graph, do –

(1) Call the starting point of the edge j, and the end point i. Then, find the $\hat{d}_i = d_j + w_{j,i}$. If $\hat{d}_i < d_i$, replace d_i by $\hat{d}_i$. It is assumed that the node stores a pointer to the previous node on the route that led to the minimum cost.

(iii) Check the possibility of further reductions in the distances. If yes, this shows that the graph has "negative cycles", and the weights will not converge. Otherwise, the algorithm is completed. This step can be omitted if all edge weights are positive.

Q.20. Give the objects of routing protocols.

Ans. The main objects of routing protocols are as follows –

(i) The protocol should be distributed.
(ii) The end-to-end transmission time should be reduced.
(iii) The protocol should be bandwidth efficient.
(iv) The lifetime of the network should be maximized.
(v) The protocol should be able to react quickly to changes in the topology or link states.
(vi) The energy consumption should be kept as small as possible.

Q.21. Discuss the source routing protocol.

Ans. Each information-producing node in source routing can specify the seqeunce of nodes that this packet must take through the network. The sequence of nodes is added to the data packet, so that each node on the route knows the node to which the packet should be transmitted next. The source routing is loop free thus, that a packet will not return to an intermediate node that it had already visited.

Link state advertisements are sent out periodically by each node to neighboring nodes in *proactive source routing*. The received link state is then compared by those nodes with the one that they have stored in their local tables. If the link state is fresher, then nodes update their table and forward the information to their neighbours, and so on. Therefore, refreshed link-state information is propagated throughout the network. As the number of nodes in the network increases, the amount of information that has to be distributed also increases rapidly. That is why, proactive source routing algorithms are not suitable for large networks, and specifically not for networks whose states change frequently.

Dynamic source routing (DSR) reduces the overhead by performing on-demand routing. The routing procedure involves two steps – an initial route discovery, and route maintenance. In route discovery, the network is flooded with *route request packets*. The route request consists of the identification (ID) of the intended information destination, a unique packet ID, and a list of nodes visited by the message. On receiving a route request packet, a node checks whether it is either the desired destination or has a path to the destination stored in its own routing table. If it is not so, the route request is rebroadcasted by the node, adding its own address to the list of nodes visited in the message. If the node is the destination, then the node answers with a route reply packet. The route reply packet tracks back along the recognized path to the source, and finally notifies the source about the sequence of nodes that have to be taken from source to destination.

During route maintenance, the protocol checks whether links in the established route are broken. Then it either uses an alternative stored route for the destination, or initiates another new route discovery process.

Q.22. Explain link-state based routing.

Ans. In link-state routing, each node collects information about the states of the links in the total network. On the basis of this information, a node can then form the most efficient routes through the network to all other nodes, for example, by using Dijkstra algorithm.

The link states are need to be acquired and distributed throughout the network in link-state based routing. The link-state information distribution to other nodes can be obtained by means of short messages. Those messages have the following information – (i) the ID of the node which is creating the advertisement (ii) the nodes to which the advertising node is attached. (iii) the sequence number which represents how fresh the information is.

The optimized link state routing protocol (OLSR) is a famous algorithm for implementing link state based routing for wireless networks. In OLSR, messages between nodes are exchanged regularly to form routing tables at every node. The classical link-state protocols flood the link-state information throughout the network, while OLSR limits this information. For this reason, OLSR is well suited for wireless networks. Hence, OLSR uses the concept of *multi-point relays* (MPRs).

Q.23. Discuss distance vector routing.

Ans. Here, each node keeps a list of all destinations that only contains the cost of getting to that destination, and the next node to send the message to. Therefore, in distance vector routing the source node only knows to which node to deliver the packet, which in turn knows the next node, and so on. The benefit of this approach is the reduced storage costs than link-state algorithms. The distance vector algorithms are thus, easier to implement and need less storage space. The actual route can be obtained using the Bellman-Ford algorithm. However, there are some drawbacks of DVR –

(i) Slow Convergence – Bellman-Ford needs multiple passes of the cost information which makes it slower than Dijkstra algorithm.

(ii) Counting to Infinity – In the extreme case that part of the network becomes separated, the network can create loops, resulting in information comes back to a node that it had already visited. The problem is such that even if node Q tells node P that it has a route to the destination, node P does not know if that route has node P. This is not a problem under normal converged situations because a route having a loop has a higher total cost compared to a route that cuts out that loop. However, in case that a node goes down, a loop can be formed. Suppose a linear network connected as P-Q-R-

S-T-U, and let the edge cost be unity for each hop. Now if node P goes down, node Q will not find an update about node P during the update process of Bellman-Ford and thus assumes that the link to node P is down. However, at the same time, node Q receives an update from node R, which tells node Q that node P is only two jumps from node R. Thus, this misinformation propagates through the network, until the nodes have increased their costs to infinity.

The counting-to-infinity problem can be solved by means of *destination sequences*, resulting in the DSDV (Destination-Sequenced Distance Vector) algorithm. For this case, the nodes store the cost of getting to the destination and the next node on the route as well as a sequence number. Then, routes are advertised periodically by each node to other destinations. The destination increases the sequence number and propagates the route to other nodes in the network. The selected route contains the largest sequence number in the network. When a link breakes, the sequence number is increased and the node cost is increased to infinity. This change is immediately propagated to other nodes in the network.

Q.24. Explain geography-based routing.

Ans. The nodes in a network mostly know their position. This can be achieved by global positioning system GPS localization. Routes can be designed on the basis of this algorithm.

Geographical routing schemes depend on the concept of "progress" toward the destination. Greedy schemes select the node that has the smallest distance to the destination within the coverage disk. Alternatively, they can maximize the projection of the link which connects the transmitting node to a specific receiving node, onto the line connecting the transmitting node to the destination. This method is known as *most forward within radius*. Fig. 5.11 shows an example to differentiate the behaviour of these two algorithms. However, mostly the two algorithms find the same path.

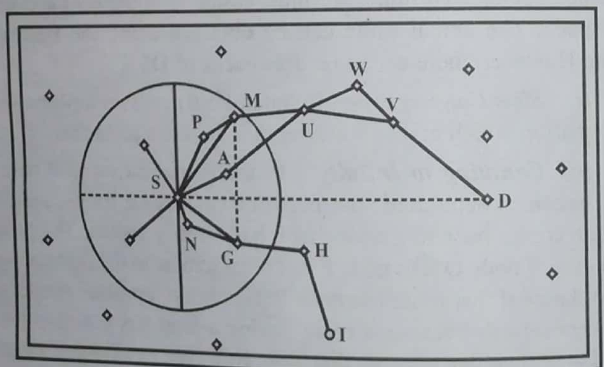


Fig. 5.11 Criteria for Choosing Nodes in Geographic Routing

Q.25. Write short note on hierarchical routing.

Ans. In this type of routing, the nodes of the network are subdivided into group of nodes, called clusters. One node in the cluster is called the *clusterhead* and the remaining nodes in the cluster can only communicate with this particular node. A message can be passed between clusters only by communications through the clusterheads. A node can only join that cluster, whose clusterhead it can talk to with minimum energy expenditure. When all nodes are battery powered, then the role of the clusterhead can be randomly rotated between the nodes in a cluster. The probability of a node to become a clusterhead is proportional to the remaining battery power of the node. When there are certain nodes that are connected to power mains, so that energy consumption is not an important matter, then those nodes are always used as clusterheads. The LEACH protocol is a popular hierarchical algorithm.

Q.26. Write down the impact of node mobility.

Ans. The nodes are static during operation in most of the ad hoc networks. However, there are also certain networks where nodes exhibit extreme mobility, for example, in vehicular ad hoc networks.

The main benefit of node mobility is that data packets can hitchhike on moving nodes. Suppose that a data packet is to be sent from node X to node Y, both of which are static, and far away. Then, node X sends the packet to a highly mobile node Z when it is passing by node X. Node Z stores the message, and when it comes close to node Y, sends it. Therefore, the large distance between nodes X and Y can be bridged without using either high transmit power for direct transmission, or multiple transmissions.

The main disadvantage of node mobility is that the network can become temporarily disconnected, particularly in sparse networks where there are only few possible routes between a source and a destination. If very less number of nodes are moving, it can be possible that there is no valid path for a multi-hop connection between source and destination anymore.

Q.27. Discuss the various power allocation strategies.

Ans. The various power allocation strategies are –

(i) Routes Fixed, Transmission Rate Fixed – In this case, we can only reduce the expended power. The transmit power at each node should be reduced as much as possible with the constraint that the SNR at the receiving node has to be high enough to ensure decoding.

(ii) Routes Fixed, but Transmission Rate Variable – In this case, if the transmit power is reduced then the necessary transmission time is increased. If the main objective of the optimization process is to minimize the energy expenditure, then transmit power should be reduced as much as possible. If the message has to be sent to the destination within a specific deadline, then an optimization of the power allocation can be done for a given delivery time.

(iii) Routes and Transmission Rates Variable – We can adapt the edge weights of the graph representing the network by changing the power and/or the rate. Therefore, an optimal route for a specific set of transmit powers might not be optimum for another set. Thus, one joint step is used for routing and power allocation.

Q.28. Discuss a stochastic network optimization algorithm such as backpressure.

Ans. The backpressure algorithm is a stochastic network optimization algorithm that turns out to be optimal under some situations. The formulation as a control problem allows to use techniques from control theory such as Lyapunov functions. In this approach each node contains a buffer in which it stores arriving data, and it tries to send data to empty the buffer.

Consider a network setup containing N nodes, connected by L links. The transmission of the messages happens in a slotted way, where t denotes the slot index. The link between two nodes x and y is characterized by a transmission rate $\mu_{xy}(t)$. The summary of rates is given in the transmission matrix $\mu(t) = \mathbf{C}(\mathbf{I}(t), S(t))$ where C denotes the transmission rate function. The transmission rate function C relies on the network topology state S(t) which describes all the effects that the network cannot influence and the link control action $\mathbf{I}(t)$ which includes all the actions of the network which can be affected, such as power control, etc. The capacity-achieving transmission is the most important example of a transmission rate function, so that on the *l*-th link.

$$C_l(\mathbf{P}(t), S(t)) = \log_2\left[1 + \frac{P_l(t)\,\alpha_{ll}(S(t))}{P_n + \sum\limits_{k\neq l} P_k(t)\,\alpha_{kl}(S(t))}\right]$$

where $\alpha_{kl}(S(t)) = |h_{kl}|^2$ denotes the power gain (inverse attenuation) of the signal transmitted to the intended receiver of link *l* by the intended transmitter of the link k.

An amount of data R_q^{out} is being sent by a node during each timeslot. At the same time, data are coming from an external source and the node is receiving data through the wireless links from other sources. The total amount of data coming during one timeslot is R_q^{in}. Each node contains an infinitely large buffer, to store the arriving messages prior to their transmission over the wireless links. The amount of data in the buffer is denoted as $Q_q(t)$ where q denotes the index of the considered queue. All backlogs are written into a

vector $\mathbf{Q}(t)$. The change of backlog of a queue at a particular node during one timeslot is given as,

$$Q_q(t+1) \leq \max\left[Q_q(t) - R_q^{out}(I(t), S(t)), 0\right] + R_q^{in}(I(t), S(t))$$

where the max[., 0] operator ensures that the queue length does not become negative.

Q.29. Write short note on routing and resource allocation in multi hop networks. ***(R.G.P.V., June 2014)***

Or

Explain the mechanism of routing and resource allocation in multi hop networks. ***(R.G.P.V., June 2015)***

Or

Discuss routing and resource allocation in multi hop networks. ***(R.G.P.V., June 2016)***

Or

Discuss about routing and resource allocation in multi hop networks. ***(R.G.P.V., June 2017)***

Ans. For larger networks, a several parallel relays are not sufficient to obtain the information from the source to the destination. Rather, we have to use multiple relays sequentially. This section assumes multi hop systems in which a packet of data can be transferred through multiple relays using multi hop decode and forward (MDF). The packet is sent from the source to the first relay node, in which this is decoded, re-encoded and transmitted to next relay, which also does decode/re-encode, then sends it to the next relay and so on. Each of the hops can be operated as a point-to-point link.

The combined routing and resource allocation is a cross layer design problem. Therefore, routing algorithms goes to consider a fixed physical layer, for which an optical route is constructed. Conversely, other papers consider a given route, and try to optimize the PHY for this route.

Also refer to the ans. of Q.18, Q.19, Q.21, Q.22 and Q.27.

ROUTING AND RESOURCE ALLOCATION IN COLLABORATIVE NETWORKS, APPLICATIONS, NETWORK CODING

Q.30. Discuss edge-disjoint routing and anypath routing in collaborative networks.

Ans. Edge-disjoint shortest-path routing is a way of determining routes which do not share any links. A minor modification of the Bellman-Ford algorithm is a suitable algorithm. However, this approach does not make much use of the broadcast effect.

Anypath routing utilizes the broadcast effect to obtain diversity. The data packet is broadcasted by each node to a group of neighbours, known as the forwarding set. The routing can continue as long as at least one of the nodes in the forwarding set receives the message. The next relay on the route is one of the successful nodes. In other terms, the broadcast effect is used to get selection diversity. Therefore, the data packet can reach its destination even if a particular link on the nominal shortest path goes down, without requiring to search for a new route. Thus, anypath routing is suitable for links that frequently change in quality.

Anypath routing results in an ensemble of possible routes in place of a specific route. A packet can take different routes through the network depending on the outputs of the transmissions from various nodes. In other words, an anypath route is the union of all possible trajectories along which a packet can travel from the source to the destination. There is a tradeoff in finding the best anypath route – increasing the candidate set gives a better robustness, but a larger forwarding set also increases the danger that the packet is routed farther away from the true shortest path. The problem becomes more complicated when the data rate is also permitted to vary. However, in any case, variations of the Bellman-Ford algorithm can obtain the optimum anypath route in polynomial time.

Q.31. Discuss routing with energy accumulation.

Ans. Energy accumulation at the relay nodes is another way of using diversity. This happens when a node stores a received signal of a packet that is very weak for decoding and combines it with another signal of the same packet which arrives later. When employing energy accumulation at the nodes instead of simple multi-hopping, the optimum route alters.

The issues involved in the problem of finding the optimum route are – (i) what nodes should participate, and (ii) in what sequence, and with what power. The former problem is NP-hard, that is, it can be solved exactly only by trying out all possible node combinations, and selecting the best one. Several heuristic algorithms are available for finding the best route. Some of those algorithms begin with the optimum multihop route, and then add on nodes that decrease the overall energy consumption. Another kind of algorithm generates the route from scratch, beginning from the source node. When it adds the next relay on the path, it decreases the signal energy still needed at all the other nodes. This reduction of energy relies on how much energy those other nodes can overhear when the new node transmits. In either case, the energy savings from the energy accumulation increase because the density of nodes increases. That is, the closer the nodes, the more energy a node can overhear.

Q.32. Explain the use of fountain codes for routing.

Ans. Relay nodes can utilize overhearing the signals intended for other relay nodes in a much more efficient manner using Fountain codes. They accumulate mutual information, instead of energy. Routing with mutual-information accumulation has two similar properties as with energy accumulation – (i) finding an optimum route is NP-hard, and (ii) for heuristic algorithms, it is better to divide the problem into two subproblems – finding the physical route or order of nodes through which packets propagate, and the allocation of resources (time, power) among the nodes. It is assumed that each node has a fixed transmission power, the optimum allocation of time can be done by a linear program for a specific routing order. Then, on the basis of the results of the LP, a simple algorithm can revise the routing order. A very efficient approach to obtain good route can be yielded by iterating between the two subproblems even in very large neworks.

The LP can be set up the following way – by the end of the k time interval, the total information flow to the k-th node from the k–1 nodes ahead of it in the route must exceed the packet payload of B bits. Mathematically,

$$\sum_{i=0}^{k-1}\sum_{n=0}^{k-1} A_{i,n}C_{i,k} \geq B$$

where $A_{i,n}$ denotes the resource allocated to transmitter i in the nth time interval and $C_{i,k}$ denotes the data rate from node i to node k. These constraints along with the objective of "minimization of total energy" form an LP that can be solved by standard software packages. The LP solution is then used to update the route. If the start time of the k + 1 time interval is similar to that of the k-th time interval, the sequence of the k-th and k + 1-th node on the route are swapped. When a relay node swaps its place with the destination in the decoding sequence, it is not used at all.

Q.33. Write short note on routing and resource allocation in collaborative networks. ***(R.G.P.V., June 2014)***

Or

Explain the mechanism of routing and resource allocation in collaborative networks. ***(R.G.P.V., June 2015)***

Or

Discuss about routing and resource allocation in collaborative networks. ***(R.G.P.V., June 2016, 2017)***

Ans. The routing problem even for a single message becomes much more complex, when collaborative communications is used for the forwarding of messages. For multi-hop routing, all we required to found was the sequence of nodes that forward the message, and the transmit power used by every node. If cooperative communications are used, start and stop time of transmission from a relay are arbitrary parameters that need to be optimized.

In contrast to the multi-hop case an optimum solution can be obtained by trying out all possible routes. In this reason, a several heuristic algorithms have been proposed that can often come close to ideal performance. In addition, it should be noted that the field of collaborative routing is away from mature. In particular, there is a dearth of solutions for the case when multiple messages are to be transmitted by the network simultaneously.

Also refer to the ans. of Q.30, Q.31 and Q.32.

Q.34. Discuss the applications of relaying and multi-hopping.

Ans. The relaying and multi-hopping can be used either in an infrastructure-based setting to facilitate the communication between a BS and an MS or they can be used as an integral part of ad hoc networks.

(i) Dedicated Relays – Mostly, dedicated relay stations (RSs) are used in cellular networks. The RS may be useful in one or more of the following respects –

(a) Increase of Coverage Area – MSs that are far away from the BS can still receive a decodable signal because the relay improves the SNR. This can increase the effective radius of a cell, or remove coverage holes, that is, offer coverage in areas within a cell that are not covered by the BS because of the peculiarities of the topography. Dedicated RSs are generally placed at locations where they have good connection to the BS for example, on rooftops. This permits the RS to get the signal with high quality and forward it to the destination. The signal arriving at the MS has improved SNR in either case.

(b) Improvement of Indoor Coverage – Relays are also useful to improve the indoor coverage. The coverage on the streets of urban/metropolitan areas is generally very good, while indoor coverage on the same streets is often spotty or nonexistent because of the additional path loss suffered by signals when they penetrate into the building. Therefore, relays are essential to fully cover the inside of office or residential buildings. Similarly, relays can also help to improve connections to the inside of trains, buses, and other vehicles.

(c) Increase of Reliability – Relays can improve the reliability by improving the SNR. In addition, relays can add diversity, which also increases reliability.

(d) Increase of Throughput – The higher throughput can be obtained if both the BS-RS and RS-MS links have good SNR. For example, by using a higher order modulation alphabet and higher coding rate can give higher throughput. However, most relays lose rate because of the half duplexing constraint. The tradeoff between the gains in per-link data rate and the duplexing loss decide whether a relay can help to increase the capacity and throughput of a cellular system or not. Another approach to increase throughput is the use of relays to redirect traffic from overloaded BSs to less-congested BSs. Consider a case where a cell is temporarily overloaded. Then some of the MSs calls can be connected through relays to BSs in neighboring cells which are momentarily less utilized.

(ii) Relaying and User Cooperation in Ad-hoc Networks – A wireless ad-hoc network is a decentralized wireless network that does not use fixed infrastructure. Depending on the requirements of particular data packets, each node can function as source, relay, and destination. It means that all nodes are created equal. The use of multiple nodes increases the reliability and reduces the costs.

Ad-hoc networks also make the basis of sensor networks. Generally, sensor networks are networks of nodes which can sense physical data and communicate them to a data sink. This data sink can be a computer-human interface or an automated control/monitoring system. Data packets have to be relayed by other nodes in the network, Similar to ad-hoc networks, because generally the sensing nodes cannot reach the data sink in a single hop. However, some of the important differences are as follows –

(a) Even though the source nodes differ the data sink is the same.

(b) The information received from the different sensing nodes is correlated.

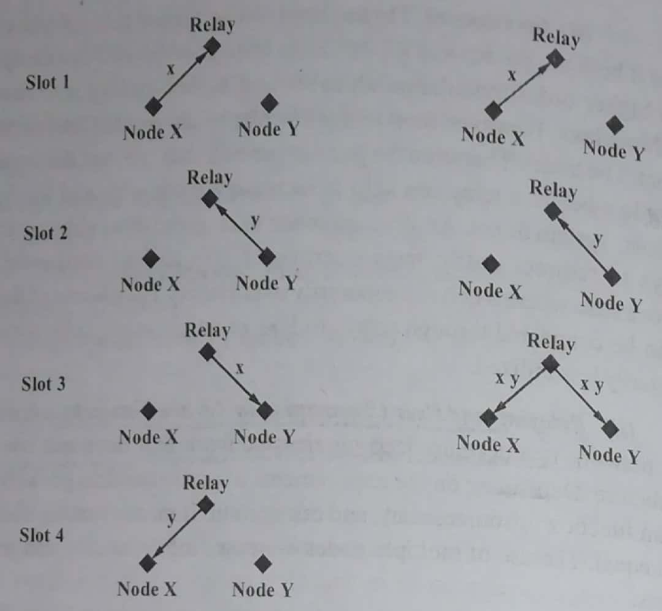
Q.35. Write short note on network coding.
(R.G.P.V., June 2014, 2015, 2017)

Or

Give a brief note on network coding. ***(R.G.P.V., June 2016)***

Ans. The basic principle of network coding is that nodes in the network make combinations of messages that they get and forward combinations of messages. Then, the destination recovers the original messages from different combinations it receives. Therefore, network coding can be considered as the epitome of collaboration.

Network coding is useful mainly for multicast situations, that is, where multiple sources transmit messages, and all nodes wish to learn those messages.



(a) Conventional Method *(b) Network Coding*

Fig. 5.12 Bidirectional Relaying

The bidirectional relay is the simplest example of a network code. Consider the situation shown in fig. 5.12. Here, two messages x and y, are need to be exchanged between nodes X and Y through relay node R. The conventional method of relaying is to use, TDMA to separate the messages, and thus needs four time slots for the exchange – (i) in slot 1, node X sends message x to the relay, (ii) in slot 2, node Y transmits message y to the relay, (iii) in slot 3, the relay sends message x to node Y, (iv) in slot 4, the relay sends message y to node X.

The more effich approach for doing the same is as follows – slots (i) and (ii) are same as before, that is, nodes X and Y separately send their messages to the relay. In the third slot, the relay broadcasts the sum of the two messages s = x + y. Node X can easily determine message y from the sum signal, y = s – x because it already knows message x. Likewise, node Y can find message x from the sum signal s. This approach improves the spectral efficiency of the transmission. Here, the sum signal s actually is a symbol-by-symbol summation of the messages, not a concatenation of the two messages and therefore has the same length as the individual message x.

In this example, we have assumed that the relay adds the complex modulation symbols of the two messages. This approach results in a higher power value of the sum signal, which is not desirable.

In an alternative approach the relay can demodulate the messages x and y, and then do a modulo-2 addition (XOR) of the information bits. Generally, this approach is used for network coding. Other approaches of network coding include the use of novel modulation constellations onto which received complex symbol combinations are mapped.

Q.36. Write down the applications of network coding to wireless systems.

Or

Write down the practical applications of network coding.

(R.G.P.V., June 2016)

Ans. The applications of network coding to wireless systems are as follows –

(i) They can be used advantageously in multicasting scenarios. However, most wireless traffic today is unicast, for which network coding is not necessarily optimal.

(ii) Most of the current network coding theory ignores noise and interference. The zero-interference assumption is fulfilled approximately if only one node that is adjacent to an intended receiving node can transmit at one time in an ad hoc network; the other adjacent nodes have to be silent. However, this approach needs additional overhead for the coordination of the transmission times and decrease the network's spectral efficiency.

Q.37. Write short note on interference alignment.

Ans. Interference alignment can be used to offer each user with half the capacity it would get in an interference-free environment, irrespective of the number of users. Using this approach the sum capacity of a network increases linearly with the number of users, which remains constant with a conventional multiple access approach. Several methods have been proposed for implementing such interference alignment. The easiest method uses time variations of the wireless propagation channel – characterize the channels between the N transmitters and the N receivers with a N×N matrix –

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} & \cdots & h_{1N} \\ h_{21} & h_{22} & h_{23} & \cdots & h_{2N} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ h_{N1} & h_{N2} & \cdots & \cdots & h_{NN} \end{bmatrix}$$

The signals transmitted when the channel has one specific realization are repeated when the channel is in its complementary state –

$$\begin{bmatrix} h_{11} & -h_{12} & -h_{13} & \cdots & -h_{1N} \\ -h_{21} & h_{22} & -h_{23} & \cdots & -h_{2N} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -h_{N1} & -h_{N2} & \cdots & \cdots & h_{NN} \end{bmatrix}$$

Then, the receivers just have to add up signals from those two transmissions, to receive effectively transmitted signals over a diagonal channel –

$$\begin{bmatrix} h_{11} & 0 & 0 & \cdots\cdots & 0 \\ 0 & h_{22} & 0 & \cdots\cdots & 0 \\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots \\ 0 & 0 & \cdots\cdots & \cdots\cdots & h_{NN} \end{bmatrix}$$

Thus, there is an interference-free signal for each user. However, the capacity is reduced by a factor 2 because of the required signal repetition.

The implementation of the concept can be done with very little hardware effort. However, it can result in large delays in the signal transmission, since the transmitter has to wait until the channel takes on the complementary state. If there are slow temporal channel variations then the delay can be longer.

OOO