CS 480: MOBILE NETWORKS

Lecture PowerPoints

SIGNAL TRANSMISSION FUNDAMENTALS

Topics Covered

- General Introduction
- Data Communication Concepts
- Basic Elements of a Communications System
- Essential Features of Communication
- How Signals Can Be Used To Convey Information
- Concept of Signals and Signal Spectrum
- Metrics of Signal Strength and Bandwidth
- Noise, Bandwidth and Channel Capacity
- Analogue and Digital Transmission

Learning Outcomes

The following are the expected learning outcomes:

- Understand waveforms representation in time and frequency domains
- Understand metrics of analogue bandwidth, power bit rate and throughput
- Understand Shannon's and Nyquist theorems

General Introduction (1)

- These lectures are intended to provide a review of transmission fundamentals
- The concepts of data communication are established and how signals can be used for conveying information identified.
- The concept of signals in the time and the frequency domain are established and various signal spectrums identified.
- The characteristics of signals and the metrics of signal strength and bandwidth are established and the relationship between data rate and bandwidth shown

General Introduction (2)

- Shannon's theorem is used to show the relationship of noise, bandwidth and channel capacity and the degenerative effect of noise on information in a system.
- The major differences between analogue and digital transmission are discussed with relation to Nyquist bandwidth.
- This subject will concentrate on the concepts and their consequences for wireless systems
- The mathematical treatment of the subject is avoided.

Data Communication Concepts

- The term "**communications**" refers to the sending, receiving and processing of information by some means.
- **Data communication** is the transmission of electronic data over some media.
  - The fundamental purpose of data communications is to exchange information between two agents.
- The hardware and software systems that enable the transmission of data make up data **communication networks**.
- Adding users to the communication networks make up a **communication system**.
- Basic building blocks of telecommunications systems
  - consist of computers, communication channel, communication equipment and communication software

Basic Elements of a Communications System (1)

- Telecommunications may be defined as
  - any process that permits the passage from sender to one or more receivers of information of any nature delivered in an easy to use form by any electromagnetic system
- Examples of easy to use form
  - printed copy, visible or audio signals etc
- electromagnetic system
  - electrical transmission by wire, radio, optical transmission, guided waves etc
- Telecommunications include telegraphy, telephony, video-telephony, data transmission, etc.
- Data communications
  - may be defined as the subset of telecommunications that involves the transmission of data to and from computers and components of computer systems.

Basic Elements of a Communications System (2)

- More specifically, data communications is
  - the transmission of data through conducted communication media such as wires, coaxial cables, or fibre optic cables, or radiated communication media such as spread spectrum radio signals, infrared light, and microwaves

Essential Features of Communication

- Data communication has several important features.
- Communication of any kind requires
  - a message,
  - a sender,
  - a receiver,
  - and a medium.
- In addition, the receiver should be able to understand the message.
- In data communication systems, some means of error detection are typically included
  - in order to ensure that the content of the message received is the same as what was sent.

Message

- For two entities to communicate, there must be a message to be exchanged between them.
- Messages can assume several forms
  - such as text, audio, video, or image and can be of varying lengths
- In business data communications, messages include
  - files,
  - requests for data/information,
  - responses to requests,
  - network status information,
  - network control messages,
  - and correspondence

Sender

- In a communication system, the sender is the transmitter of the message.
- A sender can either be a person or a machine.
- In today's networks, the sender is typically a computer
- However, in some business networks, the sender may be a terminal with enough intelligence to originate a message or response without human intervention.
- Sensors, scanners, and other input devices may also be senders in today's networks.

Receiver

- Receivers can include
  - computers, terminals, network printers, display devices, people, and computer-controllable devices such as drill presses, lightning systems, and air conditioners.
- Even though a message and a sender can exist without a receiver, communication cannot take place unless there is a receiver.
- In a computer network, a server may transmit a status message to all attached personal computers (PCs) informing them that a server will be shut down in 20 minutes to perform maintenance,
  - but if all the PCs are turned off at that moment, no communication has occurred.

Medium (1)

- Messages are carried from sender to receiver through some medium of communication.
  - In oral communications, sound waves are transmitted through air (the medium)
  - In a data communications, the term medium refers to the actual carrier of data signals between senders and receivers.
- Business data communication networks often use a variety of media to transmit data, including wires, radio waves, and light pulses.

Medium (2)

The other representation of data communication is as follows:

- Source
  - This device generates the data to be transmitted; examples are telephones and personal computers
- Transmitter
  - Usually, the data generated by a source system are not transmitted directly in the form in which they were generated.
  - Rather, a transmitter transforms and encodes the information to produce electromagnetic signals that can be transmitted across some sort of transmission system.
  - E.g., a modem takes a digital bit stream from an attached device such as a PC and transforms that bit stream into an analog signal that can be handled by telephone network.

Medium (3)

- Transmission system:
  - This can be a single transmission line or a complex network connecting source and destination.
- Receiver:
  - Accepts the signal from the transmission system and converts it into a form that can be handled by the destination device.
  - E.g., a modem accepts an analog signal coming from a network or transmission line and convert it into a digital bit stream.
- Destination:
  - Takes the incoming data from the receiver.

Simplified Data Communications Model (1)

- Figure 1 shows a simplified data communications model.
  - We trace the details of this figure using e-mail as an example.
- Suppose that the input device and transmitter are components of a personal computer (PC).
- The PC user wishes to send message *m* to another user.
- The user activates the e-mail package on the PC and enters the message via the keyboard (input device).
- The character string is briefly buffered in main memory.
- We can view it as a sequence of bits (g) in memory.
- The PC is connected to some transmission medium,
  - such as a local network or a telephone line, by an I/O device (transmitter), such as a local network transceiver or a modem.

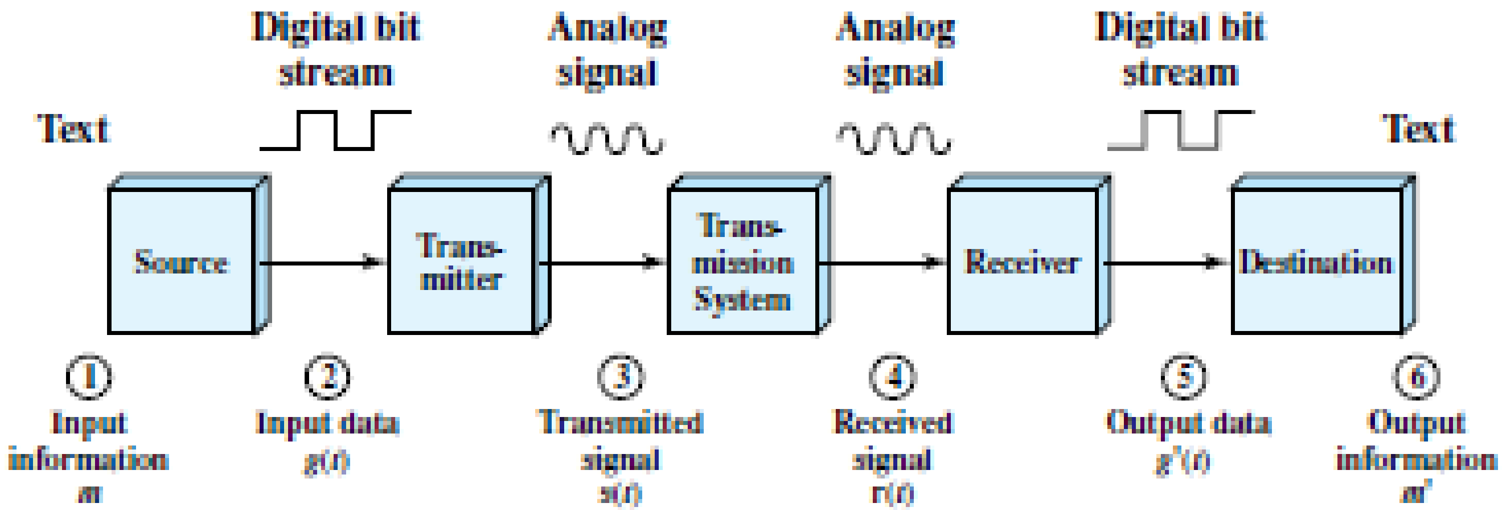


Figure 1: Simplified Data Communications Model

Simplified Data Communications Model (2)

- The input data are transferred to the transmitter as a sequence of voltage shifts [g(t)]
  - representing bits on some communications bus or cable.
- The transmitter is connected directly to the medium
  - and converts the incoming stream [g(t)] into a signal [s(t)] suitable for transmission
- The transmitted signal s(t) presented to the medium is subject to a number of impairments.
  - Thus, the received signal r(t) may differ from s(t).
- The receiver attempt to estimate the original s(t), based on r(t) and its knowledge of the medium, producing a sequence of bits g'(t).

Simplified Data Communications Model (3)

- These bits are sent to the output PC, where they are briefly buffered in memory as a block of bits (g’).
- In many cases, the destination system will attempt to determine if an error has occurred
  - and, if so, cooperate with the source system to eventually obtain a complete, error-free block of data.
- These data are then presented to the user via an output device, such as a printer or screen.

How Signals Can Be Used To Convey Information (1)

- To communicate, a device must interface with the transmission system.
- All the forms of communication depend on the use of electromagnetic signals propagated over a transmission medium
- Thus, once an interface is established, signal generation is required for communication.
- The properties of the signal, such as form and intensity, must be such that
  - the signal is capable of being propagated through the transmission system, and interpretable as data at the receiver.

How Signals Can Be Used To Convey Information (2)

- There must be some form of synchronization between transmitter and receiver
  - The receiver must be able to determine when a signal begins to arrive and when it ends.
- In all communication systems, there is potential for error
  - transmitted signals are distorted before reaching their destination
- Error detection and correction are required in circumstances where errors cannot be tolerated.
- Flow control is required to assure that the source does not overwhelm the destination by sending data faster than they can be processed and absorbed

Concept of Signals and Signal Spectrum (1)

- Signals are the physical representation of data.
- Users of a communication system can only exchange data through the transmission of signals
- The physical layer (layer 1) of the ISO/OSI basic reference model is responsible for the conversion of data
  - i.e. bits into signals and vice versa
- The physical layer is concerned with the transmission of unstructured bit stream over physical medium,
  - and deals with the mechanical, electrical, functional, as well as procedural characteristics to access the physical medium.
- Signals are functions of time and location.
- Signal parameters represent the data values.

Concept of Signals and Signal Spectrum (2)

- The most interesting types of signals for radio transmission are *periodic signals*,
  - especially *sine waves* as carriers
- The general function of a sine wave is:

$$g(t) = A_t \sin(2\pi f_t t + \varphi_t)$$

- Signal parameters are the *amplitude* A, the *frequency f*, and the phase shift ϕ.
  - The amplitude as a factor of the function g may also change over time, thus A_t
  - The frequency *f* expresses the periodicity of the signal with period $T=1/f$. The frequency may also change over time, thus f_t.
  - The phase shift determines the shift of the signal relative to the same signal without shift.

Concept of Signals and Signal Spectrum (3)

- Signal spectrum can be categorized into three namely
  - Amplitude modulation
  - Frequency modulation
  - Phase modulation

Time Domain Representation of a Signal (1)

- Figure 2 shows a sine function without a phase shift and the same function with a phase shift ϕ
- A typical way to represent signals is the time domain (see Figure 2).
  - Here the amplitude A of a signal is shown versus time.
  - Time is mostly measured in seconds, s, amplitudes can be measured in volts, V.
  - A phase shift can also be shown in this representation.
- Sine waves are of special interest
  - because it is possible to construct every periodic signal *g* by using only sine and cosine functions according to a fundamental equation of *Fourier* (see Figure 3).

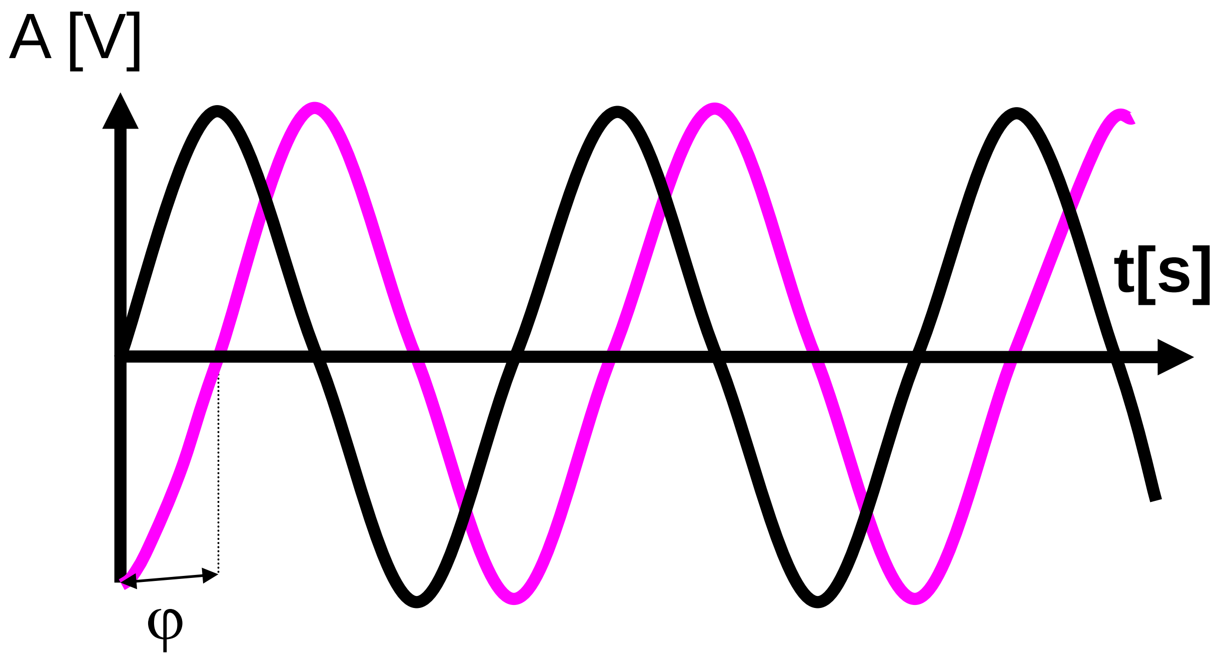
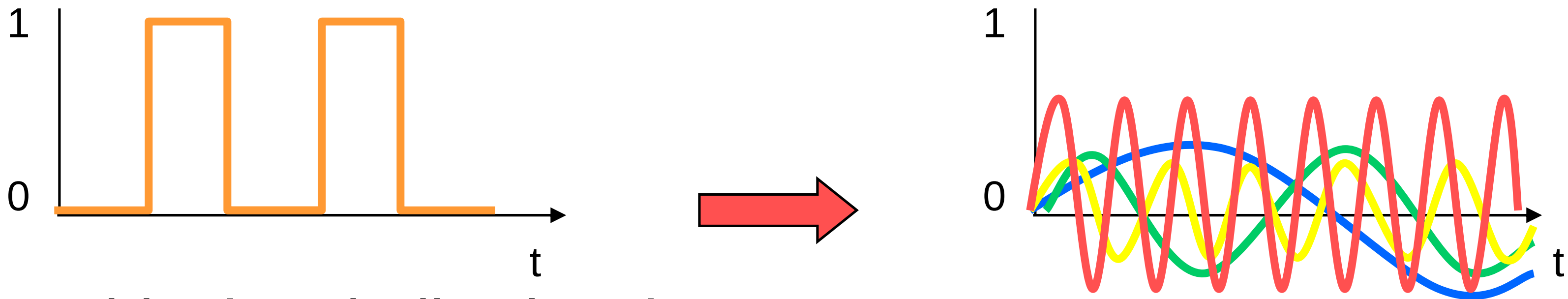


Figure 2: Time domain representation of a signal

$$g(t) = \frac{1}{2}c + \sum_{n=1}^{\infty} a_n \sin(2\pi nft) + \sum_{n=1}^{\infty} b_n \cos(2\pi nft)$$



ideal periodic signal

real composition
(based on harmonics)

Figure 3: Fundamental Fourier Equation.

Time Domain Representation of a Signal (2)

- The parameter c determines the *Direct Current* (DC) component of the signal
- The coefficients a_n and b_n are the amplitudes of the nth sine and cosine function.
- The equation shows that an infinite number of sine and cosine functions is needed to construct arbitrary periodic functions.
- The frequencies of these functions (so called harmonics) increase with increasing n and are a multiple of fundamental frequency f.

Time Domain Representation of a Signal (3)

- The bandwidth of any medium be it air, cable, transmitter, is limited and there is an upper limit for the frequencies.
- All real transmitting systems exhibit these bandwidth limits and can never transmit arbitrary periodic functions.
- It should be understood that transmitted signals are composed of one or many sine functions.

Frequency Domain Representation of a Signal

- A typical way to represent signals is the time domain
- Representations in time domain are problematic if the signal consists of many different frequencies.
- In this case, a better representation of a signal is the *frequency domain* (see Figure 4)
- Here the amplitude of a certain frequency part of the signal is shown versus frequency.
- In Figure 4, only one peak of the signal is shown.
  - The signal consists of one frequency part (a single sine function)
- Arbitrary functions would have many peaks, shown as the frequency spectrum of a signal.
- A tool to display frequencies is a spectrum analyzer.

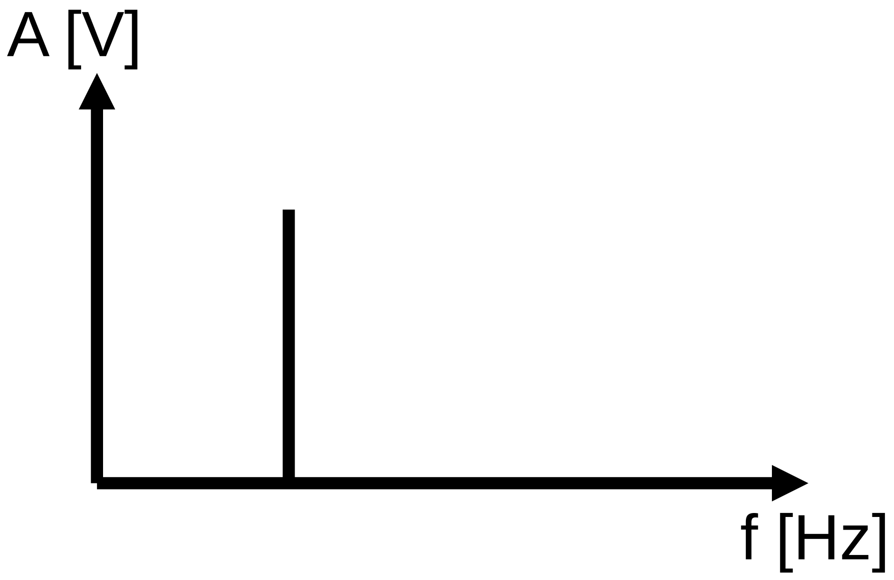


Figure 4: Frequency domain representation of a signal

Phase Domain Representation of a Signal (1)

- A third way to represent signals is the *phase domain* shown in Figure 5.
- This representation shows the magnitude M of a signal and its phase in polar coordinates.
  - It is also called phase state or signal constellation diagram
- The length of a vector represents the amplitude, the angle the phase shift.
- The x-axis represents a phase of 0^o and also called In-Phase (I)
- A phase shift of 90^o ($\pi/2$) would be a point on the y-axis, called Quadrature (Q).

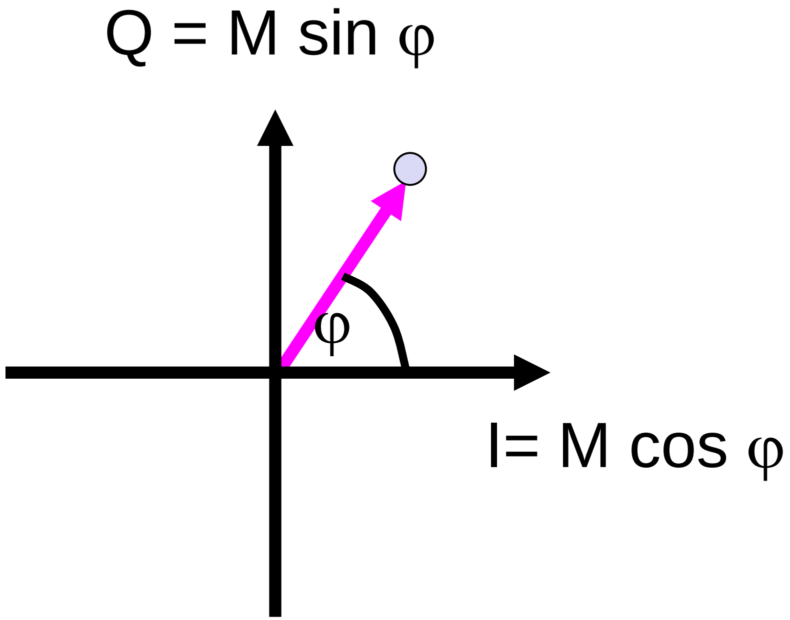


Figure 5: Phase domain representation of a signal

Metrics of Signal Strength and Bandwidth (1)

- Data transmission speeds are related to signal frequencies.
- The frequency of a signal is the number of cycles per second and is often expressed in units called **Hertz** (Hz)
  - 1Hz is 1 cycle per second.
- Various physical factors limit the maximum and minimum frequencies that a given medium can carry.
- A transmission link is called a **channel**
- The range of frequencies over which a transmission may take place over a channel is called the **bandwidth** of the channel.
- Broadband channels (i.e., those with broad bandwidth) have the ability to carry more data.

Metrics of Signal Strength and Bandwidth (2)

- **Broadband** channels are also called wideband channels
- The data-carrying capacity of a channel is of interest
- Imagine that there are two pulses per cycle
  - a “0” and a “1”
- The number of pulses per second is a unit called a **baud**
- In simple situations the baud rate is twice the frequency.
- Data can be encoded and transmitted in terms of pulses with the binary values “0” and “1”.
- However, not all the bits carry data, as in the case of stop and start bits
- The rate of data flow, called the **data transfer rate**, is therefore less than the **baud rate**.

Metrics of Signal Strength and Bandwidth (3)

- Data transfer rates are often expressed in terms of the number of data bits transmitted per second.
  - The actual units may be Kbits/s (thousands of bits per second) or Mbits/s (millions of bits per second).
- It is relatively easy to restore a **digital signal** to its original amplitude and free it of noise by means of a unit called a **digital repeater**.
- However, analog repeaters are only able to restore amplitude and, worse still, they amplify noise.
  - Therefore digital signals are preferable to analog signals.
- The bandwidth of an analog communication channel is the difference between the minimum and maximum frequencies it can carry.

Metrics of Signal Strength and Bandwidth (4)

- A voice-grade dial-up channel that can transmit frequencies between 300 and 3 400 Hz has a bandwidth of 3, 100 Hz.
- For digital circuits, bandwidth is measured in bits per second instead of a frequency range.
- For both analog and digital circuits, bandwidth is a measure of the amount of data that can be transmitted per unit of time
  - and is directly proportional to the maximum data transmission speed of a medium.
- The higher the bandwidth is, the greater the data-carrying capacity

Noise, Bandwidth and Channel Capacity

- Noise
- Bandwidth and Channel Capacity
- Channel Capacity and Shannon's Theorem

Noise

- For any data transmission event, the received signal will consist of
  - the transmitted signal, modified by the various distortions imposed by the transmission system,
  - plus additional unwanted signals that are inserted somewhere between transmission and reception.
- The latter, undesired signals are referred to as noise.
  - Noise is the major limiting factor in communications system performance.
- Noise may be divided into four categories:
  - Thermal noise or White noise
  - Crosstalk
  - Intermodulation noise
  - Impulse noise

Thermal Noise or White Noise

- Thermal noise is due to thermal agitation of electrons.
- It is present in all electronic devices and transmission media and is a function of temperature.
- White noise in telephone circuits is sometimes heard as static or hissing on the line.
- Thermal noise cannot be eliminated and therefore places an upper bound on communications system performance
- Thermal noise is significant for satellite communication
  - Because of the weakness of the signal received by satellite earth stations,

Cross Talk

- Occurs when signals from one channel distort or interfere with the signals of a different channel.
  - In telephone conversations, crosstalk appears in the form of another party's conversation being heard in the background.
- It is also present in radio frequency and multiplexed transmissions
  - when the frequency ranges are too close together.
- Crosstalk in twisted-pair wire connections occurs when
  - wire pairs interfere with each other as a result of strong signals, improper shielding, or both
- Another common example is interference between receivers and transmitters when a strong outgoing signal interferers with a weaker incoming signal.

Intermodulation Noise

- Is one special form of crosstalk, which is the result of two or more signals combining to produce a signal outside the limits of the communication channel.
- When signals at different frequencies share the same transmission medium, the result may be intermodulation noise.
- The effect of intermodulation noise is to produce signals at a frequency that is the sum or difference of the two original frequencies or multiples of those frequencies.
  - For example, the mixing of signals at frequencies f_1 and f_2 and might produce energy at the frequency $f_1 + f_2$
  - This derived signal could interfere with an intended signal at the frequency $f_1 + f_2$

Impulse Noise (1)

- All types of noise discussed so far have reasonably predictable and relatively constant magnitudes.
  - It’s possible to engineer transmission system to cope with them
- Impulse noise is non-continuous, consisting of irregular pulses or noise spikes of short duration and of relatively high amplitude.
- It is generated from a variety of causes,
  - including external electromagnetic disturbances, such as lightning, and faults and flaws in the communications system.
- It is characterized by signal spikes
  - In telephone circuits, it can be caused by switching equipment or lightning strikes
- Impulse noise is the main source of error in digital data communication.

Impulse Noise (2)

- A sharp spike of energy of 0.01 s duration would not destroy any voice data but would wash out about 560 bits of digital data being transmitted at 56 kbps.
- Figure 6 is an example of the effect of noise on a digital signal
  - Here the noise consists of a relatively modest level of thermal noise plus occasional spikes of impulse noise.
  - The digital data can be recovered from the signal by sampling the received waveform once per bit time.
  - As can be seen, the noise is occasionally sufficient to change a 1 to a 0 or a 0 to a 1.

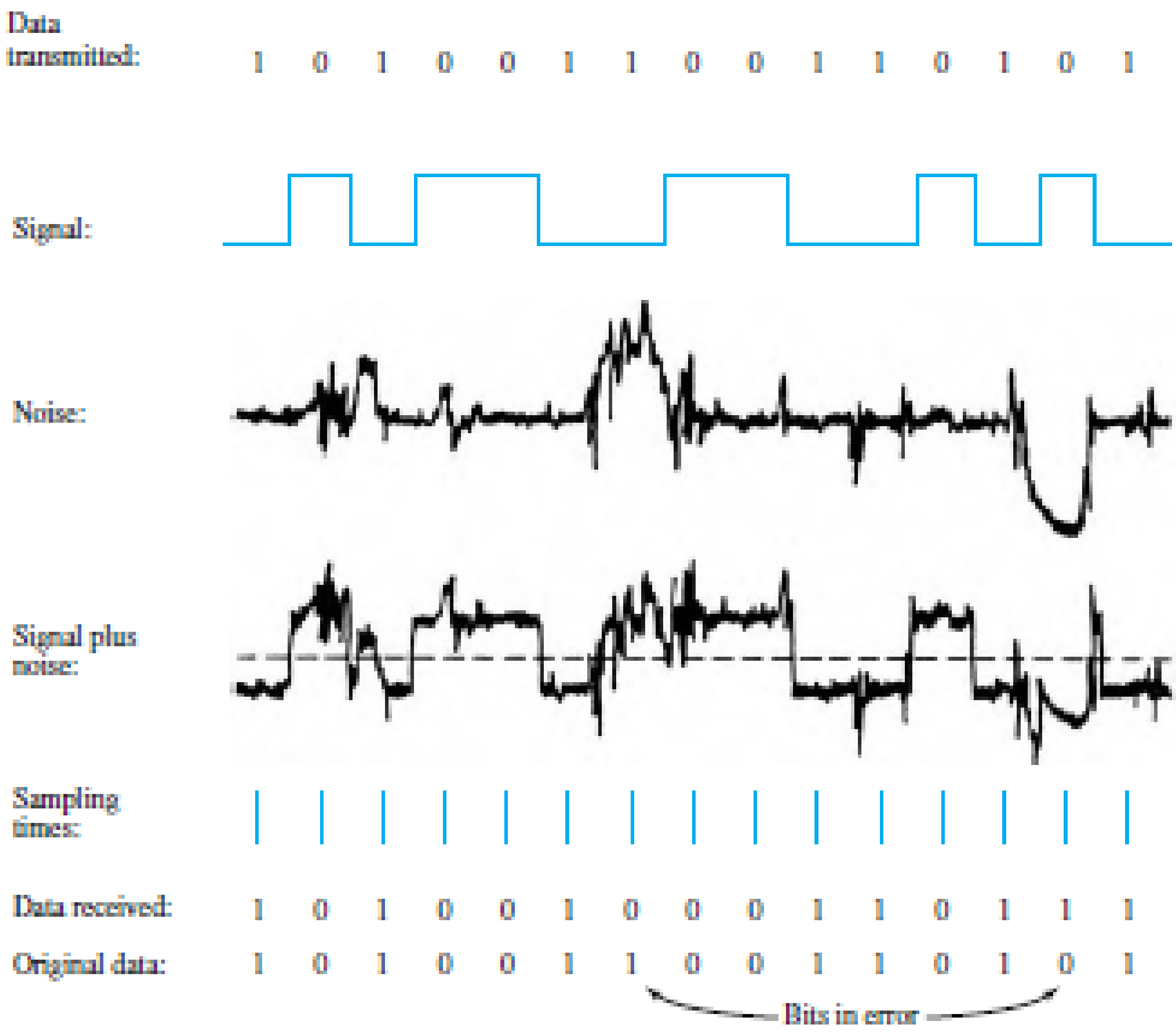


Figure 6: Effect of Noise on a Digital Signal

Bandwidth and Channel Capacity (1)

- We have seen that there are a variety of impairments that distort or corrupt a signal.
- For digital data, the question that then arises is to what extent these impairments limit the data rate that can be achieved.
- The maximum rate at which data can be transmitted over a given communication path, or channel, under given conditions, is referred to as the channel capacity.
- There are four concepts here that we are trying to relate to one another.
  - **Data rate**: The rate, in bits per second (bps), at which data can be communicated

Bandwidth and Channel Capacity (2)

- **Bandwidth**: The bandwidth of the transmitted signal as constrained by the transmitter and the nature of the transmission medium, expressed in cycles per second, or Hertz
- **Noise**: The average level of noise over the communications path
- **Error rate**: The rate at which errors occur, where an error is the reception of a 1 when a 0 was transmitted or the reception of a 0 when a 1 was transmitted

- The problem we are addressing is this:
  - Communications facilities are expensive and, in general, the greater the bandwidth of a facility, the greater the cost.
  - Furthermore, all transmission channels of any practical interest are of limited bandwidth.

Bandwidth and Channel Capacity (3)

- The limitations arise from
  - the physical properties of the transmission medium
  - or from deliberate limitations at the transmitter on the bandwidth to prevent interference from other sources.
- Accordingly, we would like to make as efficient use as possible of a given bandwidth.
- For digital data, this means that we would like to get as high a data rate as possible at a particular limit of error rate for given bandwidth.
- The main constraint on achieving this efficiency is noise.

Nyquist Bandwidth (1)

- To begin, let us consider the case of a channel that is noise free.
- In this environment, the limitation on data rate is simply the bandwidth of the signal.
- A formulation of this limitation, due to Nyquist, states that
  – if the rate of signal transmission is 2B, then a signal with frequencies no greater than B is sufficient to carry the signal rate
- The converse is also true:
  – Given a bandwidth of B, the highest signal rate that can be carried is 2B.
- This limitation is due to the effect of intersymbol interference, such as is produced by delay distortion.

Nyquist Bandwidth (2)

- In the preceding paragraph, we referred to signal rate.
  - If the signals to be transmitted are binary (two voltage levels), then the data rate that can be supported by B Hz is 2B bps.
- With multilevel signaling, the Nyquist formulation becomes

  $C = 2B\log_2 M$

  - where M is the number of discrete signal or voltage levels.
- So, for a given bandwidth, the data rate can be increased by increasing the number of different signal elements.

Nyquist Bandwidth (3)

- However, this places an increased burden on the receiver:
  - Instead of distinguishing one of two possible signal elements during each signal time, it must distinguish one of M possible signal elements.
- Noise and other impairments on the transmission line will limit the practical value of M.
- Example
  - Consider a voice channel being used, via modem, to transmit digital data. Assume a bandwidth of 3100 Hz.Then the Nyquist capacity, C, of the channel is 2B = 6200 bps. For m = 8, a value used with some modems, C becomes18,600 bps for a bandwidth of 3100 Hz.

Shannon Capacity Formula (1)

- Nyquist’s formula indicates that, all other things being equal, doubling the bandwidth doubles the data rate.
- Now consider the relationship among data rate, noise, and error rate.
- The presence of noise can corrupt one or more bits.
- If the data rate is increased, the bits become “shorter” so that more bits are affected by a given pattern of noise.
- All of these concepts can be tied together neatly in a formula developed by the mathematician Claude Shannon.

Shannon Capacity Formula (2)

- For a given level of noise, we would expect that a greater signal strength would improve the ability to receive data correctly in the presence of noise.
- The key parameter involved in this reasoning is the signal-to-noise ratio (SNR, or S/N),
  - which is the ratio of the power in a signal to the power contained in the noise that is present at a particular point in the transmission
- Typically, this ratio is measured at a receiver, because it is at this point that an attempt is made to process the signal and recover the data.

Shannon Capacity Formula (3)

- For convenience, this ratio is often reported in decibels:

$$SNR_{dB} = 10\log_{10}\left(\frac{\text{signal power}}{\text{noise power}}\right)$$

- This expresses the amount in decibels, that the intended signal exceeds the noise level.
- A high SNR will mean a high-quality signal and a low number of required intermediate repeaters.
- The signal-to-noise ratio is important in the transmission of digital data because it sets the upper bound on the achievable data rate.

Shannon Capacity Formula (4)

- Shannon's result is that the maximum channel capacity, in bits per second, obeys the equation

  $C = B \log_2(1 + SNR)$

  – where C is the capacity of the channel in bits per second and B is the bandwidth of the channel in Hertz.
- The Shannon formula represents the theoretical maximum that can be achieved.
- In practice, however, only much lower rates are achieved.
  – One reason for this is that the formula assumes white noise (thermal noise).
  – Impulse noise is not accounted for, nor are attenuation distortion or delay distortion.

Shannon Capacity Formula (6)

- Example:
  - Let us consider an example that relates the Nyquist and Shannon formulations. Suppose that the spectrum of a channel is between 3 MHz and 4 MHz and SNR_{dB} = 24 dB. Then

    $B = 4\text{MHz} – 3\text{ MHz} = 1\text{ MHz}.$

    $SNR_{dB} = 24\text{ dB} = 10 \log_{10} (SNR)$

    $SNR = 251$

    Using Shannon's formula,

    $C = 10^6 \times \log_2 (1 + 251) \approx 10^6 \times 8 = 8\text{ Mbps}$
  - This is a theoretical limit and is unlikely to be reached.
  - But assume we can achieve the limit. Based on Nyquist's formula, how many signaling levels are required? We have

    $C = 2B\log_2 M$

    $8 \times 10^6 = 2 \times 10^6 \times \log_2 M$

    $M = 16$

Analog and Digital Transmission (1)

- Signals are electric or electromagnetic encoding of data.
- Transmission is the communication of data by the propagation and processing of signals.
- The terms analog and digital correspond to continuous and discrete.
- An analog signal is a continuously varying electromagnetic wave that may be propagated over a variety of media, depending on spectrum.
- A digital signal is a sequence of voltage pulses that may be transmitted over a wire medium,
  - for example, a constant positive voltage level may represent binary 1 and a constant negative voltage level may represent binary 0.

Analog and Digital Transmission (2)

- Analog transmission
  - The transmission of signals without regard to content where the signal may be amplified, but there is no intermediate attempt to recover the data from the signal
- Digital transmission
  - The transmission of digital data, using either an analog or digital signal, in which the digital data are recovered and repeated at intermediate points to reduce the effects of noise.
- Data is transmitted over communication channels as either analog signals or digital signals.
- Analog transmission is used on traditional telephone lines:
  - Messages are sent and received over telephone wires in the form of continuous electronic waves that follow the path of the wire.

Review Questions

- Differentiate between guided media and unguided media.
- Differentiate between an analog and a digital electromagnetic signal.
- What are three important characteristics of a periodic signal?
- What is the relationship between the wavelength and frequency of a sine wave?
- What is the relationship between a signal's spectrum and its bandwidth?