Principles of Data Communications

Reference Book: Data Communications and Networking by Behrouz A. Forouzan

PERFORMANCE OF THE NETWORK

- Bandwidth
- Throughput
- Latency (Delay)
- Bandwidth-Delay Product
- Jitter

Bandwidth

- Bandwidth in Hertz
  - Bandwidth in hertz is the range of frequencies contained in a composite signal or the range of frequencies a channel can pass.
- Bandwidth in Bits per Seconds
  - The term bandwidth can also refer to the number of bits per second that a channel, a link, or even a network can transmit.
- There is an explicit relationship between the bandwidth in hertz and bandwidth in bits per second.
- Basically, an increase in bandwidth in hertz means an increase in bandwidth in bits per second.
- The relationship depends on whether we have baseband transmission or transmission with modulation.

Throughput

- The throughput is a measure of how fast we can actually send data through a network.
- A link may have a bandwidth of B bps, but we can only send T bps through this link with T always less than B.
- In other words, the bandwidth is a potential measurement of a link; the throughput is an actual measurement of how fast we can send data.
- For example, we may have a link with a bandwidth of 1 Mbps, but the devices connected to the end of the link may handle only 200 kbps. This means that we cannot send more than 200 kbps through this link.
- Imagine a highway designed to transmit 1000 cars per minute from one point to another. However, if there is congestion on the road, this figure may be reduced to 100 cars per minute. The bandwidth is 1000 cars per minute; the throughput is 100 cars per minute.

Latency (Delay)

- The latency or delay defines how long it takes for an entire message to completely arrive at the destination from the time the first bit is sent out from the source.
- We can say that latency is made of four components: propagation time, transmission time, queuing time and processing delay.
- Latency $=$ Propagation time $+$ Transmission time $+$ Queuing time $+$ Processing delay

Propagation Time

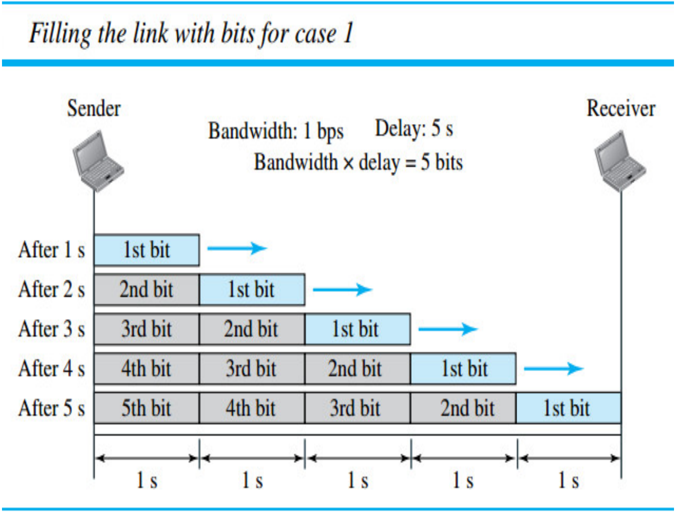
- Propagation time measures the time required for a bit to travel from the source to the destination. The propagation time is calculated by dividing the distance by the propagation speed.
- Propagation time = Distance/(Propagation Speed)

Transmission Time

- In data communications we don't send just 1 bit, we send a message.
- The first bit may take a time equal to the propagation time to reach its destination; the last bit also may take the same amount of time.
- However, there is a time between the first bit leaving the sender and the last bit arriving at the receiver. The first bit leaves earlier and arrives earlier; the last bit leaves later and arrives later.
- The transmission time of a message depends on the size of the message and the bandwidth of the channel.
- Transmission time = (Message size) / Bandwidth

Queuing Time

- The third component in latency is the queuing time, the time needed for each intermediate or end device to hold the message before it can be processed.
- The queuing time is not a fixed factor; it changes with the load imposed on the network.
- When there is heavy traffic on the network, the queuing time increases.
- An intermediate device, such as a router, queues the arrived messages and processes them one by one. If there are many messages, each message will have to wait.

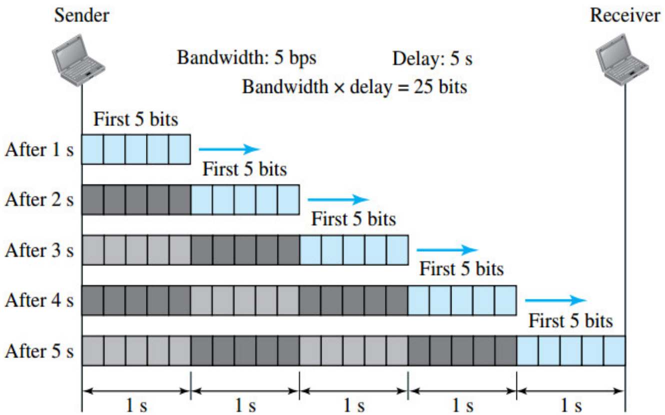

Filling the link with bits for case 1
Sender
Receiver
Bandwidth: 1 bps
Delay: 5 s
Bandwidth × delay = 5 bits
After 1 s
After 2 s
After 3 s
After 4 s
After 5 s
1st bit
2nd bit
3rd bit
4th bit
5th bit
1 s

Bandwidth-Delay Product

- Case 1. Let us assume that we have a link with a bandwidth of 1 bps (unrealistic, but good for demonstration purposes).
- We also assume that the delay of the link is 5s (also unrealistic). We want to see what the bandwidth-delay product means in this case.
- Looking at the figure, we can say that this product 1×5 is the maximum number of bits that can fill the link. There can be no more than 5 bits at any time on the link.

Bandwidth-Delay Product

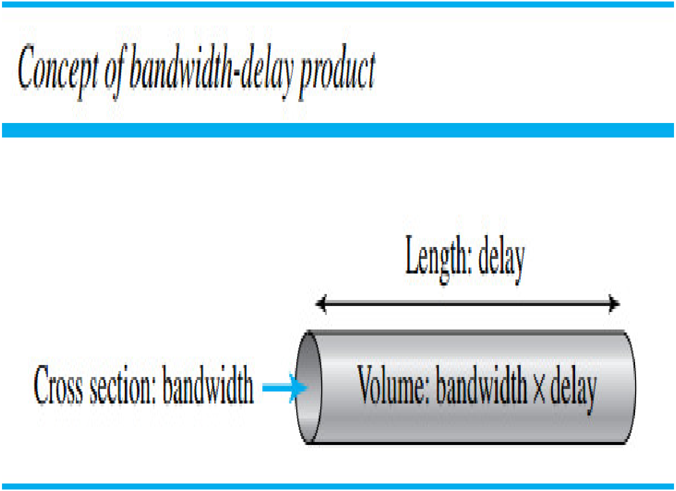
Filling the link with bits in case 2



Bandwidth-Delay Product

- Case 2. Now assume we have a bandwidth of 5 bps.
- Figure shows that there can be maximum $5 \times 5 = 25$ bits on the line.
- The reason is that, at each second, there are 5 bits on the line; the duration of each bit is 0.20 s.

Bandwidth-Delay Product



Bandwidth-Delay Product

- The bandwidth-delay product defines the number of bits that can fill the link.
- We can think about the link between two points as a pipe. The cross section of the pipe represents the bandwidth, and the length of the pipe represents the delay. We can say the volume of the pipe defines the bandwidth-delay product.

Jitter

- Another performance issue that is related to delay is jitter.
- We can roughly say that jitter is a problem if different packets of data encounter different delays and the application using the data at the receiver site is time-sensitive (audio and video data, for example).
- If the delay for the first packet is 20 ms, for the second is 45 ms, and for the third is 40 ms, then the real-time application that uses the packets endures jitter.

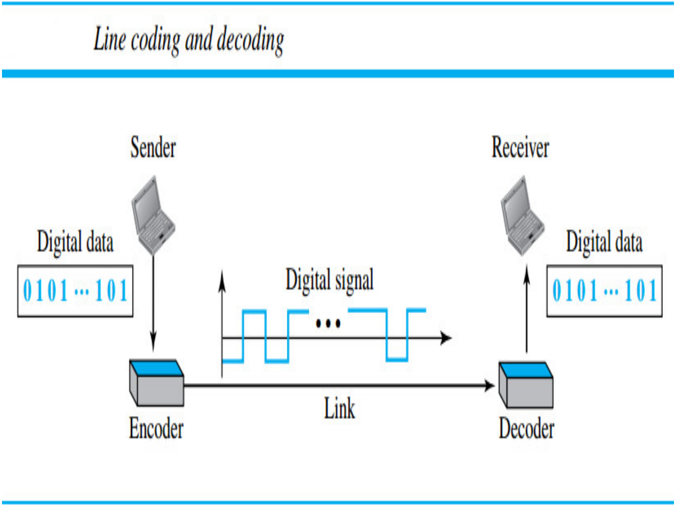
DIGITAL TRANSMISSION

- Digital-to-digital conversion techniques : methods which convert digital data to digital signals.
- Analog-to-digital conversion techniques : methods which change an analog signal to a digital signal.
- Transmission modes

DIGITAL-TO-DIGITAL CONVERSION

- The conversion involves three techniques.
  - Line coding
  - Block coding
  - Scrambling
- Line coding is always needed; block coding and scrambling may or may not be needed.

Line Coding



Line coding and decoding
Sender
Receiver
Digital data
0101 ⋯ 101
Digital signal
Digital data
0101 ⋯ 101
Encoder
Link
Decoder

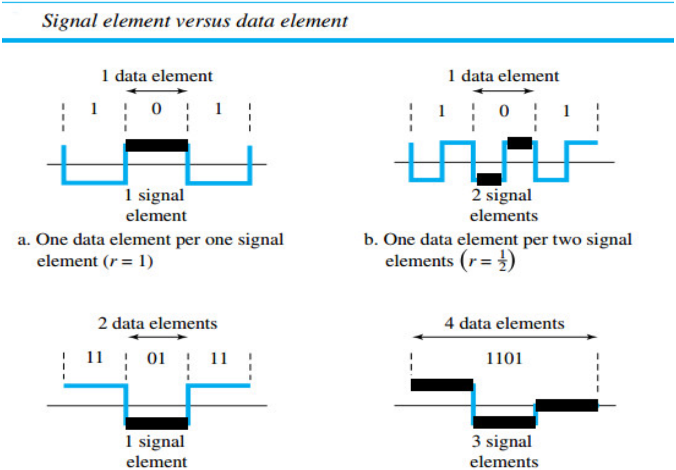
Line Coding

- Line coding is the process of converting digital data to digital signals.
- We assume that data, in the form of text, numbers, graphical images, audio, or video, are stored in computer memory as sequences of bits.
- Line coding converts a sequence of bits to a digital signal.
- At the sender, digital data are encoded into a digital signal; at the receiver, the digital data are recreated by decoding the digital signal.

Characteristics- Signal Element Versus Data Element

- Let us distinguish between a data element and a signal element.
- In data communications, our goal is to send data elements.
- A data element is the smallest entity that can represent a piece of information: this is the bit.
- In digital data communications, a signal element carries data elements.
- A signal element is the shortest unit (timewise) of a digital signal. In other words, data elements are what we need to send; signal elements are what we can send.
- Data elements are being carried; signal elements are the carriers. We define a ratio r which is the number of data elements carried by each signal element.

Characteristics- Signal Element Versus Data Element



a. One data element per one signal element ($r = 1$)

b. One data element per two signal elements ($r = \frac{1}{2}$)

c. Two data elements per one signal element ($r = 2$)

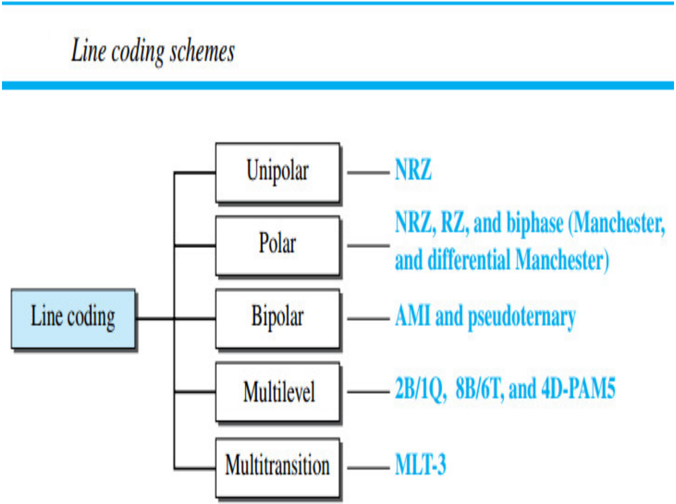
d. Four data elements per three signal elements ($r = \frac{4}{3}$)

- In part a of the figure, one data element is carried by one signal element ($r = 1$).
- In part b of the figure, we need two signal elements (two transitions) to carry each data element ($r = 1/2$).
- In part c of the figure, a signal element carries two data elements ($r = 2$). Finally, in part d, a group of 4 bits is being carried by a group of three signal elements ($r = 4/3$).
- An analogy may help here. Suppose each data element is a person who needs to be carried from one place to another. We can think of a signal element as a vehicle that can carry people.
- When $r = 1$, it means each person is driving a vehicle.
- When $r > 1$, it means more than one person is travelling in a vehicle (a carpool, for example).

Data Rate Versus Signal Rate

- The data rate defines the number of data elements (bits) sent in 1s. The unit is bits per second (bps).
- The signal rate is the number of signal elements sent in 1s. The unit is the baud.
- The data rate is sometimes called the bit rate; the signal rate is sometimes called the pulse rate, the modulation rate, or the baud rate.
- One goal in data communications is to increase the data rate while decreasing the signal rate. Increasing the data rate increases the speed of transmission; decreasing the signal rate decreases the bandwidth requirement.

Line Coding Schemes



Unipolar Scheme

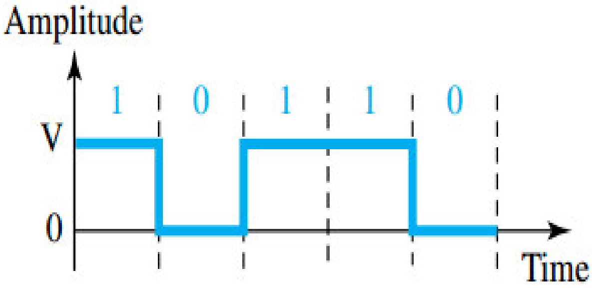
- In a unipolar scheme, all the signal levels are on one side of the time axis, either above or below.

NRZ (Non-Return-to-Zero)

- Traditionally, a unipolar scheme was designed as a non-return-to-zero (NRZ) scheme in which the positive voltage defines bit 1 and the zero voltage defines bit 0.
- It is called NRZ because the signal does not return to zero at the middle of the bit. Figure shows a unipolar NRZ scheme.
- Compared with its polar counterpart, this scheme is very costly.

NRZ (Non-Return-to-Zero)

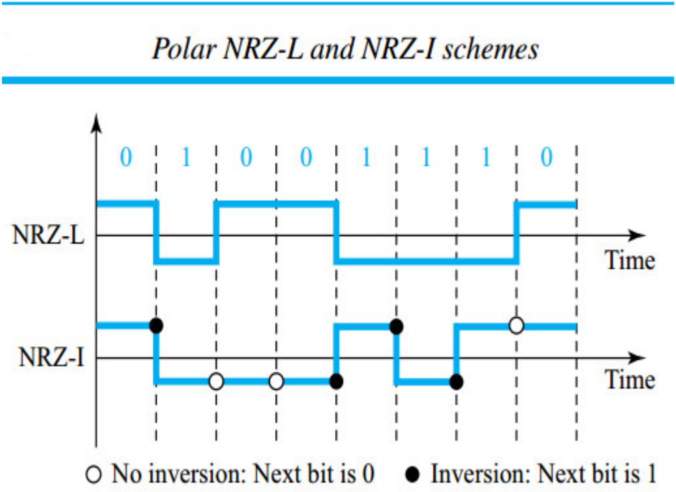
Unipolar NRZ scheme



Polar Schemes

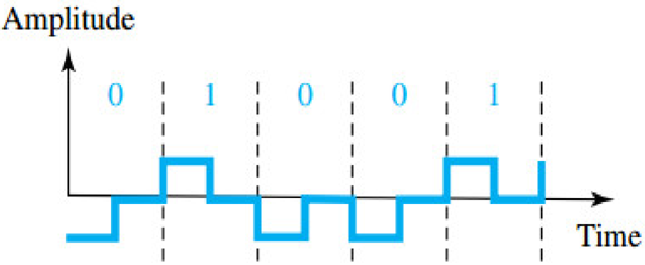
- In polar schemes, the voltages are on both sides of the time axis. For example, the voltage level for 0 can be positive and the voltage level for 1 can be negative.
- Non-Return-to-Zero (NRZ)
  - In polar NRZ encoding, we use two levels of voltage amplitude. We can have two versions of polar NRZ: NRZ-L and NRZ-I
  - In the first variation, NRZ-L (NRZ-Level), the level of the voltage determines the value of the bit.
  - In the second variation, NRZ-I (NRZ-Invert), the change or lack of change in the level of the voltage determines the value of the bit. If there is no change, the bit is 0; if there is a change, the bit is 1.

Polar Schemes



- Return-to-Zero (RZ)
  - The main problem with NRZ encoding occurs when the sender and receiver clocks are not synchronized. The receiver does not know when one bit has ended and the next bit is starting.
  - One solution is the return-to-zero (RZ) scheme, which uses three values: positive, negative, and zero.
  - In RZ, the signal changes not between bits but during the bit.
  - In Figure we see that the signal goes to 0 in the middle of each bit. It remains there until the beginning of the next bit.
  - The main disadvantage of RZ encoding is that it requires two signal changes to encode a bit and therefore occupies greater bandwidth.
  - RZ uses three levels of voltage, which is more complex to create and discern.
  - As a result of all these deficiencies, the scheme is not used today.

Polar RZ scheme

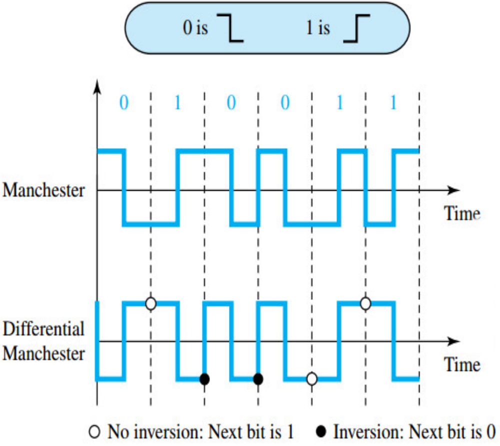


Biphase: Manchester and Differential Manchester

- The idea of RZ (transition at the middle of the bit) and the idea of NRZ-L are combined into the Manchester scheme.
- In Manchester encoding, the duration of the bit is divided into two halves.
- The voltage remains at one level during the first half and moves to the other level in the second half.
- The transition at the middle of the bit provides synchronization.
- Differential Manchester, on the other hand, combines the ideas of RZ and NRZ-I.
- There is always a transition at the middle of the bit, but the bit values are determined at the beginning of the bit. If the next bit is 0, there is a transition; if the next bit is 1, there is none.
- The Manchester scheme overcomes several problems associated with NRZ-L, and differential Manchester overcomes several problems associated with NRZ-I.

- Manchester and differential Manchester schemes are also called biphase schemes.
- **In Manchester and differential Manchester encoding, the transition at the middle of the bit is used for synchronization.**

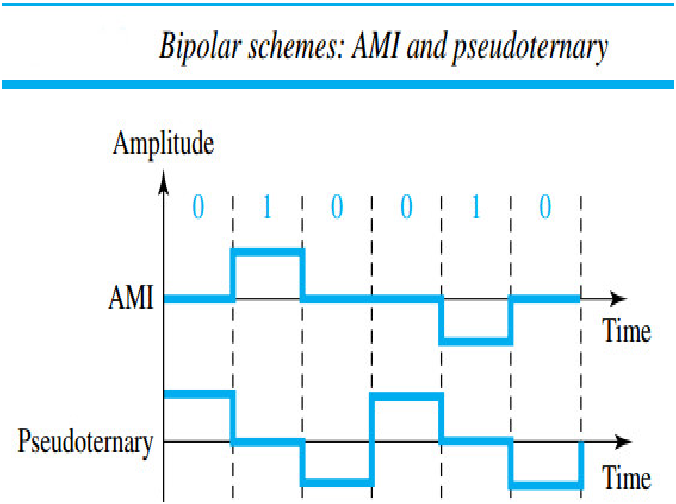
Polar biphase: Manchester and differential Manchester schemes



Bipolar Schemes

- In bipolar encoding (sometimes called multilevel binary), there are three voltage levels: positive, negative, and zero.
- The voltage level for one data element is at zero, while the voltage level for the other element alternates between positive and negative.

AMI and Pseudoternary



AMI and Pseudoternary

- Two variations of bipolar encoding: AMI and pseudoternary.
- A common bipolar encoding scheme is called bipolar alternate mark inversion (AMI).
- In the term alternate mark inversion, the word mark comes from telegraphy and means 1.
- So AMI means alternate 1 inversion.
- A neutral zero voltage represents binary 0. Binary 1s are represented by alternating positive and negative voltages.
- A variation of AMI encoding is called pseudoternary in which the 1 bit is encoded as a zero voltage and the 0 bit is encoded as alternating positive and negative voltages.

Summary

- Performance of the Network
- Digital-Digital Conversion
- Line Coding Schemes

THANK YOU