Direct Link Networks - Encoding

Reading: Peterson and Davie, Chapter 2
Where are we?

Today

- Application Layers
  - BitTorrent (P2P)
  - HTTP (Web)
  - Skype (VOIP)
  - IPTV (streaming media)

- Transport
  - TCP
  - UDP

- Network
  - IP

- Data Link
  - Ethernet
  - 4G
  - DSL
  - WiFi

- Physical

Spring 2017
© CS 438 Staff, University of Illinois
Direct Link Networks

- All hosts are directly connected by a physical medium

Key points
- Encoding and Modulation
- Framing
- Error Detection
- Medium Access Control
Internet Protocols

<table>
<thead>
<tr>
<th>Encoding</th>
<th>Framing, error detection, medium access control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
<td></td>
</tr>
<tr>
<td>Presentation</td>
<td></td>
</tr>
<tr>
<td>Session</td>
<td></td>
</tr>
<tr>
<td>Transport</td>
<td></td>
</tr>
<tr>
<td>Network</td>
<td></td>
</tr>
<tr>
<td>Data Link</td>
<td></td>
</tr>
<tr>
<td>Physical</td>
<td></td>
</tr>
</tbody>
</table>

- User-level software
- Kernel software (device driver)
- Hardware (network adapter)
Direct Link Networks - Outline

- Hardware building blocks
- Encoding
- Framing
- Error detection
- Multiple access media (MAC examples)
- Network adapters
Hardware Building Blocks

- Nodes
  - Hosts: general purpose computers
  - Switches: typically special purpose hardware
  - Routers: varied
Nodes: Workstation Architecture

- Finite memory
  - Scarce resource
- Runs at memory speeds, NOT processor speeds
Hardware Building Blocks

- Links
  - Physical medium
    - Copper wire with electronic signaling
    - Glass fiber with optical signaling
    - Wireless with electromagnetic (radio, infrared, microwave) signaling
    - Two cups and a string
Links - Copper

- Copper-based Media
  - Category 5/6 Twisted Pair
    - 10-1Gbps
    - 100m
  - ThinNet Coaxial Cable
    - 10-100Mbps
    - 200m
  - ThickNet Coaxial Cable
    - 10-100Mbps
    - 500m

- More twists, less crosstalk, better signal over longer distances

- twisted pair
- coaxial cable (coax)
  - copper core
  - insulation
  - braided outer conductor
  - outer insulation

- More expensive than twisted pair
- High bandwidth and excellent noise immunity
Links - Optical

- Optical Media
  - Multimode Fiber 100Gbps 2km
  - Single Mode Fiber 100-2400Mbps 40km
Links - Optical

- **Single mode fiber**
  - Expensive to drive (Lasers)
  - Lower attenuation (longer distances) ≤ 0.5 dB/km
  - Lower dispersion (higher data rates)

- **Multimode fiber**
  - Cheap to drive (LED’s)
  - Higher attenuation
  - Easier to terminate

\[ n_{\text{core single mode fiber}} \]
\[ \approx \text{1 wavelength thick} = \approx 1 \text{ micron} \]

\[ n_{\text{core multimode fiber}} \]
\[ \approx O(100 \text{ microns}) \text{ thick} \]
Links - Optical

- Advantages of optical communication
  - Higher bandwidths
  - Superior (lower) attenuation properties
  - Immune from electromagnetic interference
  - No crosstalk between fibers
  - Thin, lightweight, and cheap (the fiber, not the optical-electrical interfaces)
Links - Wireless

- **Path loss**
  - Signal attenuation as a function of distance
  - Signal-to-noise ratio (SNR—Signal Power/Noise Power) decreases, make signal unrecoverable

- **Multipath propagation**
  - Signal reflects off surfaces, effectively causing self-interference

- **Internal interference (from other users)**
  - Hosts within range of each other collide with one another’s transmission

- **External interference**
  - Microwave is turned on and blocks your signal
Wireless Path Loss

- Signal power attenuates by about $\sim r^2$ factor for omni-directional antennas in free space
  - $r$ is the distance between the sender and the receiver
- The exponent in the factor is different depending on placement of antennas
  - Less than 2 for directional antennas
  - Faster attenuation
    - Exponent $> 2$ when antennas are placed on the ground
    - Signal bounces off the ground and reduces the power of the signal
Wireless Multipath Effects

- Signals bounce off surfaces and interfere with one another.
- What if signals are out of phase?
  - Orthogonal signals cancel each other and nothing is received!
What is a Wireless “Link”?
What is a Wireless “Link”?
Wireless Bit Errors

- The lower the SNR (Signal/Noise) the higher the Bit Error Rate (BER)
- How can we deal with this?
  - Make the signal stronger
- Why is this not always a good idea?
  - Increased signal strength requires more power
  - Increases the interference range of the sender, so you interfere with more nodes around you
- Error correction can correct some problems
Encoding

Problems with signal transmission

- Attenuation: Signal power absorbed by medium
- Dispersion: A discrete signal spreads in space
- Noise: Random background “signals”
How can two hosts communicate?

- Encode information on modulated “Carrier signal”
  - Phase, frequency, and/or amplitude modulation
  - Ethernet: self-clocking Manchester coding
  - Technologies: copper, optical, wireless
Encoding

- **Goal**
  - Understand how to connect nodes in such a way that bits can be transmitted from one node to another

- **Idea**
  - The physical medium is used to propagate signals
    - Modulate electromagnetic waves
    - Vary voltage, frequency, wavelength
  - Data is encoded in the signal
Bauds and Bits

- **Baud rate**
  - Number of *physical symbols* transmitted per second

- **Bit rate**
  - Actual number of *data bits* transmitted per second

- **Relationship**
  - Depends on the number of *bits* encoded in each *symbol*
Analog vs. Digital Transmission

- **Analog** and **digital** correspond roughly to **continuous** and **discrete**

- **Data**: entities that convey meaning
  - **Analog**: continuously varying patterns of intensity (e.g., voice and video)
  - **Digital**: discrete values (e.g., integers, ASCII text)

- **Signals**: electric or electromagnetic encoding of data
  - **Analog**: continuously varying electromagnetic wave
    - May be propagated over a variety of media
  - **Digital**: sequence of voltage pulses
    - May be transmitted over a wire medium
Analog vs. Digital Transmission

- Advantages of digital transmission over analog
  - Cheaper
  - Simpler for multiplexing distinct data types (audio, video, e-mail, etc.)
  - Easier to encrypt

- Two examples based on modulator-demodulators (modems)
  - Electronic Industries Association (EIA) standard: RS-232
  - International Telecommunications Union (ITU) V.32 9600 bps modem standard
RS-232

- Communication between computer and modem
- Uses two voltage levels (+15V, -15V), a binary voltage encoding
- Data rate limited to 19.2 kbps (RS-232-C); raised to 115,200 kbps in later standards

Characteristics
- Serial
  - One signaling wire, one bit at a time
- Asynchronous
  - Line can be idle, clock generated from data
- Character-based
  - Send data in 7- or 8-bit characters
RS-232 Timing Diagram

One bit per clock tick

Voltage never returns to 0V

0V is a dead/disconnected line

-15V is both “idle” and “1”
RS-232

- Initiate send by
  - Push to 15V for one clock (start bit)

- Minimum delay between character transmissions
  - Idle for one clock at -15V (stop bit)

- One character
  - 0, 1 or 2 voltage transitions

- Total Bits
  - 9 bits for 7 bits of data (78% efficient)

- Start and stop bits also provide framing
RS-232 Timing Diagram

![RS-232 Timing Diagram](image-url)
Voltage Encoding

- Binary voltage encoding
  - Done with RS-232 example
  - Generalize before continuing with V.32 (not a binary voltage encoding)

- Common binary voltage encodings
  - Non-return to zero (NRZ)
  - NRZ inverted (NRZI)
  - Manchester (used by IEEE 802.3—10 Mbps Ethernet)
  - 4B/5B
Non-Return to Zero (NRZ)

- **Signal to Data**
  - High $\Rightarrow$ 1
  - Low $\Rightarrow$ 0

- **Comments**
  - Transitions maintain clock synchronization
  - Long strings of 0s confused with no signal
  - Long strings of 1s causes baseline wander
  - Both inhibit clock recovery

<table>
<thead>
<tr>
<th>Bits</th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>1</th>
<th>0</th>
<th>1</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>0</th>
</tr>
</thead>
<tbody>
<tr>
<td>NRZ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Non-Return to Zero Inverted (NRZI)

- **Signal to Data**
  - Transition $\mapsto 1$
  - Maintain $\mapsto 0$

- **Comments**
  - Solves series of 1s, but not 0s

```
Bits:  0  0  1  0  1  1  1  1  0  1  0  0  0  0  1  0

NRZ:

NRZI:
```
Manchester Encoding

- Signal to Data
  - XOR NRZ data with clock
  - High to low transition $\Rightarrow 1$
  - Low to high transition $\Rightarrow 0$

- Comments
  - (used by IEEE 802.3—10 Mbps Ethernet)
  - Solves clock recovery problem
  - Only 50% efficient ( ½ bit per transition)
4B/5B

- **Signal to Data**
  - Encode every 4 consecutive bits as a 5 bit symbol

- **Symbols**
  - At most 1 leading 0
  - At most 2 trailing 0s
  - Never more than 3 consecutive 0s
  - Transmit with NRZI

- **Comments**
  - 16 of 32 possible codes used for data
  - At least two transitions for each code
  - 80% efficient
4B/5B – Data Symbols

At most 1 leading 0
- 0000 \(\rightarrow\) 11110
- 0001 \(\rightarrow\) 01001
- 0010 \(\rightarrow\) 10100
- 0011 \(\rightarrow\) 10101
- 0100 \(\rightarrow\) 01010
- 0101 \(\rightarrow\) 01011
- 0110 \(\rightarrow\) 01110
- 0111 \(\rightarrow\) 01111

At most 2 trailing 0s
- 1000 \(\rightarrow\) 10010
- 1001 \(\rightarrow\) 10011
- 1010 \(\rightarrow\) 10110
- 1011 \(\rightarrow\) 10111
- 1100 \(\rightarrow\) 11010
- 1101 \(\rightarrow\) 11011
- 1110 \(\rightarrow\) 11100
- 1111 \(\rightarrow\) 11101

At most 1 leading 0
- 0000 \(\rightarrow\) 11110
- 0001 \(\rightarrow\) 01001
- 0010 \(\rightarrow\) 10100
- 0011 \(\rightarrow\) 10101
- 0100 \(\rightarrow\) 01010
- 0101 \(\rightarrow\) 01011
- 0110 \(\rightarrow\) 01110
- 0111 \(\rightarrow\) 01111

At most 2 trailing 0s
- 1000 \(\rightarrow\) 10010
- 1001 \(\rightarrow\) 10011
- 1010 \(\rightarrow\) 10110
- 1011 \(\rightarrow\) 10111
- 1100 \(\rightarrow\) 11010
- 1101 \(\rightarrow\) 11011
- 1110 \(\rightarrow\) 11100
- 1111 \(\rightarrow\) 11101
4B/5B – Control Symbols

- 11111 ⇒ idle
- 11000 ⇒ start of stream 1
- 10001 ⇒ start of stream 2
- 01101 ⇒ end of stream 1
- 00111 ⇒ end of stream 2
- 00100 ⇒ transmit error
- Other ⇒ invalid
Binary Voltage Encodings

- Problem with binary voltage (square wave) encodings
  - Very wide (Infinite) frequency range required, implying
    - Significant dispersion
    - Uneven attenuation
  - Prefer to use a narrower frequency band
Binary Voltage Encodings

- Problem with binary voltage (square wave) encodings
  - Very wide (Infinite) frequency range required, implying
    - Significant dispersion
    - Uneven attenuation
  - Prefer to use a narrower frequency band

- Types of modulation
  - Amplitude (AM)
  - Frequency (FM)
  - Phase/phase shift
  - Combinations of these
Example: AM/FM for continuous signal

- Original signal
- Amplitude modulation
- Frequency modulation
Amplitude Modulation

idle 1 0
Frequency Modulation

idle  1  0
Phase Modulation

idle 1 0
Phase Modulation

Phase shift in carrier frequency

108° difference in phase

collapse for 108° shift
Phase Modulation Algorithm

- Send carrier frequency for one period
  - Perform phase shift
  - Shift value encodes symbol
    - Value in range $[0, 360^\circ)$
    - Multiple values for multiple symbols
    - Represent as circle

8-symbol example
Communication between modems

Analog phone line

Uses a combination of amplitude and phase modulation

- Known as Quadrature Amplitude Modulation (QAM)

Sends one of 16 signals each clock cycle
Constellation Pattern for V.32 QAM

16-symbol example

- Same algorithm as phase modulation
- Can also change signal amplitude
- 2-dimensional representation
  - Angle is phase shift
  - Radial distance is new amplitude
Example constellation
Comments on V.32

- V.32 transmits at 2400 baud
  - *i.e.*, 2,400 symbols per second
- Each symbol contains
  - $\log_2 16 = 4$ bits
- Data rate
  - $4 \times 2400 = 9600$ bps
- Points in constellation diagram
  - Chosen to maximize error detection
  - Process called trellis coding
Modulation (Baud) Rate

- A stream of binary 1s at 1 Mbps
  - NRZI
    - 1 bit = 1 µsec
    - 1 signal element = 1 µsec
  - Manchester
    - 1 bit = 1 µsec
    - 1 signal element = 0.5 µsec

- What is a bit?
- What is a signal element?
Modulation (Baud) Rate

A stream of binary 1s at 1 Mbps

NRZI

What is the data rate?

Data Rate (R) = bits/sec = 1 Mbps for both Manchester

What is the modulation rate?

Modulation Rate = Baud Rate = Rate at which signal elements are generated = R (NRZI) = 2R (Manchester)

1 bit = 1 µsec
1 signal element = 0.5 µsec

5 bits = 5 µsec
Sampling

- Suppose you have the following 1Hz signal being received
- How fast to sample, to capture the signal?
Sampling a 1 Hz signal at 2 Hz is enough
  - Captures every peak and trough
Sampling

- Sampling a 1 Hz signal at 3 Hz is also enough
  - In fact, more than enough samples to capture variation in signal
Sampling a 1 Hz signal at 1.5 Hz is not enough

- Why?
Sampling

- Sampling a 1 Hz signal at 1.5 Hz is not enough
  - Cant distinguish between multiple possible signals
  - Problem known as **aliasing**
What about more complex signals?

- Fourier’s theorem
  - Any continuous signal can be decomposed into a sum of sines and cosines at different frequencies
- Example: Sum of 1 Hz, 2 Hz, and 3 Hz sines
  - How fast to sample?
What about more complex signals?

- Fourier’s theorem
  - Any continuous signal can be decomposed into a sum of sines and cosines at different frequencies
- Example: Sum of 1 Hz, 2 Hz, and 3 Hz sines
  - How fast to sample? --> **answer: 6 Hz**
Generalizing the Examples

- What limits baud rate?
- What data rate can a channel sustain?
- How is data rate related to bandwidth?
- How does noise affect these bounds?
- What else can limit maximum data rate?
What Limits Baud Rate?

- Baud rate
  - Typically limited by electrical signaling properties
- Changing voltages takes time
  - No matter how small the voltage or how short the wire
- Electronics
  - Slow compared to optics
- Baud rate
  - Can be as high as twice the bandwidth of communication
What Data Rate can a Channel Sustain? How is Data Rate Related to Bandwidth?

- Transmitting $N$ distinct signals over a noiseless channel with bandwidth $B$, we can achieve at most a data rate of $2B \log_2 N$.

  - **Nyquist’s Sampling Theorem** (H. Nyquist, 1920’s)
What Data Rate can a Channel Sustain? How is Data Rate Related to Bandwidth?

- Transmitting $N$ distinct signals over a noiseless channel with bandwidth $B$, we can achieve at most a data rate of

<table>
<thead>
<tr>
<th>Baud rate</th>
<th>Number of physical symbols transmitted per second</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bit rate</td>
<td>Actual number of data bits transmitted per second</td>
</tr>
</tbody>
</table>

Nyquist, 1920

Relationship
Depends on the number of bits encoded in each symbol
Noiseless Capacity

- Nyquist’s theorem: $2B \log_2 N$
- Example 1: sampling rate of a phone line
  - $B = 4000$ Hz
  - $2B = 8000$ samples/sec.
    - sample every 125 microseconds
Noiseless Capacity

- Nyquist’s theorem: $2B \log_2 N$
- Example 2: noiseless capacity
  - $B = 1200$ Hz
  - $N = \text{each pulse encodes 16 symbols}$
  - $C =$
Noiseless Capacity

- Nyquist’s theorem: $2B \log_2 N$
- Example 2: noiseless capacity
  - $B = 1200$ Hz
  - $N = \text{each pulse encodes 16 symbols}$
  - $C = 2B \log_2 (N) = D \times \log_2 (N)
    = 2400 \times 4 = 9600$ bps
How does Noise affect these Bounds?

- **Noise**
  - Blurs the symbols, reducing the number of symbols that can be reliably distinguished

- **Claude Shannon (1948)**
  - Extended Nyquist’s work to channels with additive white Gaussian noise (a good model for thermal noise)
    \[
    \text{channel capacity } C = B \log_2 (1 + S/N)
    \]

where

- \(C\) is the maximum supportable bit rate
- \(B\) is the channel bandwidth
- \(S/N\) is the ratio between signal power and in-band noise power

This \(N\) is noise, not number of symbols
How does Noise affect these Bounds?

- **Noise**
  - Blurs the symbols, reducing the number of symbols that can be reliably distinguished

- **Claude Shannon (1948)**
  - Extended Nyquist’s work to channels with additive white Gaussian noise (a good model for thermal noise)
    - \[ \text{channel capacity } C = B \log_2 (1 + S/N) \]
  - Represents error free capacity
    - also used to calculate the noise that can be tolerated to achieve a certain rate through a channel
  - Result is based on many assumptions
    - Formula assumes white noise (thermal noise)
    - Impulse noise is not accounted for
    - Various types of distortion are also not accounted for
Noisy Capacity

- Telephone channel
  - 3400 Hz at 40 dB SNR

Decibels (dB) is a logarithmic unit of measurement that expresses the magnitude of a physical quantity (usually power or intensity) relative to a specified or implied reference level.
Decibels

- A ratio between signal powers is expressed in decibels
  \[ \text{decibels (db)} = 10 \log_{10}(P_1 / P_2) \]
- Used in many contexts
  - The loss of a wireless channel
  - The gain of an amplifier
- Note that dB is a relative value
  - Can be made absolute by picking a reference point
    - Decibel-Watt – power relative to 1W
    - Decibel-milliwatt – power relative to 1 milliwatt
Signal-to-Noise Ratio

- Signal-to-noise ratio (SNR, or S/N)
  - Ratio of the power in a signal to the power contained in the noise
  - Typically measured at a receiver

- A high SNR
  - High-quality signal

- Low SNR
  - May be hard to “extract” the signal from the noise

- SNR sets upper bound on achievable data rate

\[
(SNR)_{dB} = 10 \log_{10} \left( \frac{\text{signal power}}{\text{noise power}} \right)
\]
Noisy Capacity

- Telephone channel
  - 3400 Hz at 40 dB SNR
  - \( C = B \log_2 (1 + S/N) \) bits/s
  - SNR = 40 dB
    - \( 40 = 10 \log_{10} \left( \frac{P_{\text{signal}}}{P_{\text{noise}}} \right) \)
    - \( S/N = 10,000 \)
  - \( C = 3400 \log_2 (10001) = 44.8 \) kbps

\[ \text{SNR(dB)} = 10 \log_{10} \left( \frac{P_{\text{signal}}}{P_{\text{noise}}} \right) \]
More examples of Nyquist and Shannon Formulas

- Spectrum of a channel between 3 MHz and 4 MHz; $\text{SNR}_{\text{dB}} = 24$ dB
  \[ B = \]

  $\text{SNR} = \]

- Using Shannon’s formula
  \[ C = B \log_2 (1 + \frac{S}{N}) \]
More examples of Nyquist and Shannon Formulas

- Spectrum of a channel between 3 MHz and 4 MHz; $\text{SNR}_{dB} = 24 \text{ dB}$
  $$B = 4 \text{ MHz} - 3 \text{ MHz} = 1 \text{ MHz}$$
  $$\text{SNR}_{dB} = 24 \text{ dB} = 10 \log_{10} \left( \frac{\text{SNR}}{\text{SNR}_{dB}} \right)$$
  $$\text{SNR} = 251$$

- Using Shannon’s formula
  $$C = B \log_2 \left( 1 + \frac{S}{N} \right)$$
  $$C = 10^6 \times \log_2 \left( 1 + 251 \right) \approx 10^6 \times 8 = 8 \text{ Mbps}$$
More examples of Nyquist and Shannon Formulas

How many signaling levels are required?

\[ C = 2B \log_2 M \]
More examples of Nyquist and Shannon Formulas

- How many signaling levels are required?
  \[ C = 2B \log_2 M \]
  \[ 8 \times 10^6 = 2 \times (10^6) \times \log_2 M \]
  \[ 4 = \log_2 M \]
  \[ M = 16 \]

- Look out for: dB versus linear values, \( \log_2 \) versus \( \log_{10} \)
Summary of Encoding

- Problems: attenuation, dispersion, noise
- Digital transmission allows periodic regeneration
- Variety of binary voltage encodings
  - High frequency components limit to short range
  - More voltage levels provide higher data rate
- Modulation schemes
  - Amplitude, frequency, phase, and combinations
  - Quadrature amplitude modulation: amplitude and phase, many signals
- Nyquist (noiseless) and Shannon (noisy) limits on data rates