Fundamentals of Information Transmission and Coding (a.k.a. Communication Theory)

Signals and functions

Elementary operation of communication: send signal on medium from point A to point B.

- media—copper wire, optical fiber, air/space, etc.
- signals—voltage and currents, light pulses, radio waves, microwaves, etc.

Signal can be viewed as a time-varying function $s(t)$. 
If $s(t)$ is “sufficiently nice” (Dirichlet conditions), then $s(t)$ can be represented as a linear combination of complex sinusoids.
Fourier expansion and transform:

\[ s(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S(\omega) e^{i\omega t} \, d\omega, \]

\[ S(\omega) = \int_{-\infty}^{\infty} s(t) e^{-i\omega t} \, dt. \]

\[ \rightarrow \text{time domain vs. frequency domain} \]
Example (square wave):
Random function (i.e., white noise) has “flat-looking” spectrum.

\[
\rightarrow \text{ unbounded bandwidth}
\]

Luckily, most “interesting” functions arising in practice are far from random; in fact, \textit{bandwidth limited}.

E.g., speech: 20 Hz–20 kHz; telephone systems: 300 Hz–3300 Hz

\[
\rightarrow \text{ bandwidth 3000 Hz}
\]
Digital data vs. analog data

*Digital data*: bits.

\[ \rightarrow \text{discrete signal (both in time and amplitude)} \]

*Analog data*: audio/voice, video/image; analog data is oftentimes *digitized* so that bits form the starting point.

\[ \rightarrow \text{continuous signal} \]

**Sampling theorem (Nyquist)**: Given continuous bandlimited signal $s(t)$ with $S(\omega) = 0$ for $|\omega| > W$, $s(t)$ can be reconstructed from its samples if

\[ \nu > 2W \]

where $\nu$ is the sampling rate.
\[ \nu_1 = \frac{1}{T_1} \quad < \quad \nu_2 = \frac{1}{T_2} \]
Information transmission over noiseless channel

Set-up:

- source $s$ emits symbols from finite alphabet set $\Sigma$
- symbol $a \in \Sigma$ is generated with probability $p_a > 0$
- a code book $F$ of $\Sigma$ assigns code word $w_a = F(a)$ for each $a \in \Sigma$
- $F$ is invertible
- $|w_a|$ denotes length of $w_a$ in bits
- average code length $L$ associated with $\langle \Sigma, p, F \rangle$

$$L = \sum_{a \in \Sigma} p_a |w_a|$$

Question: Given $\langle \Sigma, p \rangle$, are there “good” code books $F$ with small $L$?
Answer: Yes.

The entropy $H$ of $\langle \Sigma, p \rangle$ is defined as

$$H = \sum_{a \in \Sigma} p_a \log \frac{1}{p_a}$$

**Source-coding Theorem (Shannon):** For all $F$,

$$H \leq L.$$ 

Moreover, there exist $F$ such that

$$L \leq H + \epsilon$$

where $0 < \epsilon < 1$.

$\rightarrow$ holds for extension codes (blocks of $n$ symbols)
Information transmission over noisy channel

Channel capacity $C$: maximum achievable “reliable” data transmission rate (bps) over a noisy channel.

Channel Coding Theorem (Shannon): Given signal power $P_S$, noise power $P_N$, and channel subject to white Gaussian noise (detailed conditions omitted),

$$C = \frac{1}{2} \nu \log(1 + \frac{P_S}{P_N}) \text{ bps.}$$

Here $P_S/P_N$: signal-to-noise ratio.
By the Sampling Theorem,

\[ C = W \log(1 + \frac{P_S}{P_N}) \text{ bps}. \]

Signal-to-noise ratio (SNR) is expressed as

\[ \text{dB} = 10 \log_{10}(\frac{P_S}{P_N}). \]

**Example:** Assuming a decibel level of 30, what is the channel capacity of a telephone line?

**Answer:** First, \( W = 3000 \text{ Hz}, \frac{P_S}{P_N} = 1000 \). Using Channel Coding Theorem,

\[ C = 3000 \log 1001 \approx 30 \text{ kbps}. \]

Compare against 28.8 kbps modems.
Digital vs. analog signals

In essence, \textit{square wave} versus everything else.

Two forms of \textit{transmission}:

\begin{itemize}
  \item digital transmission: data transmission using square waves
  \item analog transmission: data transmission using all other waves
\end{itemize}

Analog data via analog transmission is straightforward; e.g., telephone at home.

Need to consider:

\begin{itemize}
  \item analog data via digital transmission
  \item digital data via analog transmission
  \item digital data via digital transmission
\end{itemize}
First, why is digital transmission “superior” to analog transmission?

Common to both: problem of *attenuation*.

- decrease in signal strength as a function of distance
- increase in attenuation as a function of frequency

Rejuvenation of signal via amplifiers (analog) and repeaters (digital).
Delay distortion: different frequency components (in guided media) travel at different speeds.

Most problematic: effect of noise (thermal, cross talk, etc.)

- Analog: Amplification also amplifies noise—filtering out just noise, in general, is a hard problem; e.g., voice.

- Digital: Repeater just generates a new square wave. Noise cannot be confused with data (unless too high).
Analog transmission of digital data

Three pieces of information to manipulate: amplitude, frequency, phase.

- Amplitude modulation (AM)—sensitive to power fluctuations.
- Frequency modulation (FM)—allow full duplex communication; need four frequencies.
- Phase modulation (PM)—more sophisticated.
Baud rate vs. bit rate

*Baud rate*: Unit of time within which carrier wave can be altered for AM, FM, or PM.

At every signalling event, potentially more than 1 bit of information can be encoded; e.g., PM with multiple phases, say, four.

Thus 2 bits of information per signalling period.

\[ \rightarrow \text{bit rate (bps)} = 2 \times \text{baud rate} \]

Combine the three; e.g., QAM—8 phase angles and 2 amplitudes for a total of 16 detectable events (CCITT v.29 standard, 9600 bps, 2400 baud).

\[ \rightarrow 4 \text{ bits per baud} \]
**Broadband vs. baseband**

Presence/absence of *carrier wave*; allows many channels to co-exist at the same time.

\[\rightarrow\] frequency division multiplexing (FDM)

Clearly, BW of medium $> 4 \times$ BW of signal.
In the absence of carrier wave, can still use multiplexing:

\[\rightarrow \text{ time-division multiplexing (TDM)}\]

Clearly, bit rate of medium > data rate of signal; however, mostly used for digital transmission of digital or analog data (PCM, codec).
Example: T1 carrier.

Assuming 4 kHz telephone channel bandwidth, Sampling Theorem dictates 8000 samples per second (125 μsec/sample).

Bandwidth = 8000 × 193 = 1.544 Mbps
Digital transmission of digital data

Direct encoding of square waves using voltage differentials; e.g., -15V—+15V for RS-232-C.

NRZ-L (non-return to zero, level), NRZI (NRZ invert on ones);

Manchester (biphase or self-clocking codes).
Trade-offs:

- NRZ codes—long sequences of 0’s (or 1’s) causes synchronization problem; need extra control line (clock) or sensitive signalling equipment.

- Biphase codes—synchronization easily achieved through self-clocking; however, other things being equal, achieves only 50% efficiency vis-a-vis NRZ codes.

4B/5B code

Encode 4 bits of data using 5 bit code where the code word has at most one leading 0 and two trailing 0’s.

0000 ↔ 11110, 0001 ↔ 01001, etc.

→ using 4B/5B, at most three consecutive 0’s

→ efficiency: 80%
Multiplexing techniques:

- TDM
- FDM
- mixture (FDM + TDM); e.g., TDMA (time division multiple access) scheme in wireless media
- spread spectrum or CDMA (code division multiple access); competing scheme with TDMA for wireless media

*Code division multiplexing*

Direct sequence:

To send (i.e., encode) bit sequence \( x = x_1, x_2, \ldots, x_n \), use pseudorandom bit sequence \( y = y_1, y_2, \ldots, y_n \) to compute

\[
\begin{align*}
z &= z_1, z_2, \ldots, z_n \\
&= x_1 \oplus y_1, x_2 \oplus y_2, \ldots, x_n \oplus y_n.
\end{align*}
\]
To decode bit sequence $z = z_1, z_2, \ldots, z_n$, compute

$$x = z \oplus y.$$ 

- data rate usually slower than code rate (spreading)
- multiplexing $N$ sources achieved via a set of chipping codes

$$\{y^1, y^2, \ldots, y^N\}$$

Frequency hopping:

Use pseudorandom number sequence as key to index a set of carrier frequencies $f_1, f_2, \ldots, f_m$ (spreading).

Receiver with access to pseudorandom sequence can decode transmitted signal.

$\longrightarrow$ code narrowband input as broadband output
Some benefits:

- more secure (eavesdropping)
- resistant to jamming (esp. freq. hopping)
- noise resistant (esp. direct sequence)
- graceful multiplexing degradation
Synchronous vs. asynchronous transmission

→ framing problem

Asynchronous: e.g., ASCII character transmission between dumb terminal and host computer.

```
1 |   |   |   |   |   |   | 0
```

stop bit                      start bit

Each character is an independent unit; receiver needs to know bit duration.

Overhead problem; assuming 1 start bit, 1 stop bit, 8 data bits, only 80% efficiency.

→ inefficient for long messages
Synchronous: “Byte-oriented scheme”; e.g., BISYNC

<table>
<thead>
<tr>
<th>SYN</th>
<th>SYN</th>
<th>SOH</th>
<th>Header</th>
<th>STX</th>
<th>Body</th>
<th>ETX</th>
<th>CRC</th>
</tr>
</thead>
</table>

→ SYN, SOH, STX, ETX, DLE: sentinels

Two problems:

- How to maintain synchronization if |Body| is large?
- Control characters within Body of message.

→ inefficient for short messages

→ efficiency approaches 1 as |Body| → ∞
“Bit-oriented scheme”; e.g., HDLC

Use fixed *preamble* and *postamble*; simply a bit pattern.

\[ \rightarrow \ 01111110 \]

How to avoid confusing 01111110 in the data part?

\[ \rightarrow \ \text{bit stuffing} \]
SONET (Synchronous Optical Network)

→ framing/transmission standard for optical fiber

Rates: STS-1 (51.84 Mbps), STS-3 (155.25 Mbps), STS-3c, STS-12c (622.08 Mbps), STS-24c (1.24416 Gbps), STS-48c, etc.

Common to use OC-\(n\) in place of STS-\(n\).

STS-1 frame:
Features:

- 125 $\mu$s frame duration (for all STS-$n$)
- 51.84 Mbps = 810 $\cdot$ 8 $\cdot$ 8000
- 3 + 1 columns of overhead
- overhead includes synchronization, pointer fields
- overhead encoded using NRZ
- payload scrambled (XOR’ed) to achieve approximate self-clocking
- SONET also used for FDDI
STS-3c frame:

- contiguous payload area—SPE (synchronous payload envelop)
- STS-3c frame can carry about 44 ATM cells
- most relevant frame (with STS-12c) for ATM networking
Error-detection and correction

General theory: subject of Information Theory.

E.g., Hamming codes, Huffman codes, Shannon-Fano codes, Reed-Solomon, etc.

Intuitive idea: Want to transmit 8-bit words reliably; use, e.g., 12-bit code words.
In network protocol context: want practical error detection.

\[ \rightarrow \text{error-correction: use retransmission} \]
\[ \rightarrow \text{two-level scheme} \]

*Parity*: Odd or even parity; single bit error detection.

*(Internet) Checksum*: Group message into 16-bit words; calculate their sum (one’s complement); append “checksum” to message.

*Cyclic redundancy check* (CRC): Polynomial arithmetic over finite field.

View \( n \)-bit string \( a_{n-1}a_{n-2}\cdots a_0 \) as a polynomial of degree \( n - 1 \):

\[
M(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1x + a_0.
\]
Fix some generator polynomial $G(x)$ of degree $k$. Choice of $G(x)$ is important.

Let $R(x)$ be the remainder of $x^k M(x) / G(x)$.

Let $T(x) = x^k M(x) - R(x)$.

$\longrightarrow T(x)$ is the code word

Assume $T(x) + E(x)$ arrives at the receiver.

If $E(x) = 0$ then $T(X)/G(x) = 0$.

$\longrightarrow$ no errors

If $E(x) \neq 0$ then $T(X)/G(x) \neq 0$.

$\longrightarrow$ error has occurred
Specific instances:

If $E(x) = x^i$, $0 \leq i \leq n + k - 1$ (i.e., a single error at position $i$), then assuming $G(x)$ contains at least two terms, $G(x)$ will fail to divide $E(x)$.

If $E(x) = x^i + x^j$, $i > j$, then first express $E(x) = x^j(x^{i-j} + 1)$.

Assuming $x$ does not divide $G(x)$, to detect double errors it is sufficient that $G(x)$ not divide $x^{i-j} + 1$.

Fact: $G(x) = x^{15} + x^{14} + 1$ will not divide $x^r + 1$ for all $k < 32768$. 
Some commonly used CRC generator polynomials:

- **CRC-32**: \( x^{32} + x^{26} + x^{23} + x^{22} + x^{16} + x^{12} + x^{11} + x^{10} + x^8 + x^7 + x^5 + x^4 + x^2 + x + 1 \) (FDDI, Ethernet)

- **CRC-CCITT**: \( x^{16} + x^{12} + x^{5} + 1 \) (HDLC)

- **CRC-8**: \( x^8 + x^2 + x + 1 \) (ATM)