Chapter 1 Introduction

The study of communication systems can be divided into two distinct areas:

1. How communication systems work.
2. How they perform in the presence of noise.

The study of each of these two areas, in turn, requires specific tools. To study the first area, the students must be familiar with signal analysis (Fourier techniques), and to study the second area, a basic understanding of probability theory and random processes is essential.

Definition: Communications is the transfer of information at one time or location to another time or location.

1. Communication system:

A typical communication system can be modeled as shown in Fig 1.1. The components of a communication system are as follows:

- The source originates a message, such as a human voice, a television picture, or data. If the data is nonelectrical (human voice, television picture), it must be converted by an input transducer into an electrical waveform referred to as the baseband signal or message signal.
- The transmitter modifies the baseband signal for efficient transmission.
- The channel is a medium - such as wire, coaxial cable, or a radio link - through which the transmitter output is sent.
- The receiver reprocesses the signal received from the channel.
- The destination is the unit to which the message is communicated.

2. Analog and digital messages:

Messages are digital or analog.

* Digital messages are constructed with a finite number of symbols. For example, a Morse-coded telegraph message is a digital message constructed from a set of only two symbols - mark and space. It is therefore a binary message, implying only two symbols. A digital message constructed with M symbols is called an M-ary message.
* Analog messages, on the other hand, are characterized by data whose values vary over a continuous range. For example, a speech waveform has amplitudes that vary over a continuous range.
* Digital versus Analog communications:
  - Digital Communications System
    - transmit a finite number of signals
    - text and data are naturally digital information sources
  - Analog Communications
    - transmit a continuous (uncountably infinite) range of signals
    - voice and video are natural analog information sources
  - Analog information source can be converted into a digital source by
    - Sampling the signal in time
    - Quantizing the signal amplitude to a finite number of levels

3. Signal-to-noise ratio, channel bandwidth, and the rate of communication

The fundamental parameters that control the rate and quality of information transmission are the channel bandwidth B and the signal power S.

The bandwidth of a channel is the range of frequencies that it can transmit with reasonable fidelity. For example, if a channel can transmit with reasonable fidelity a signal whose frequency components occupy a range from 0 (dc) up to a maximum 5 kHz, the channel bandwidth is 5 kHz.

Generally, if a channel of bandwidth B can transmit N pulses per second, then to transmit KN pulses per second we need a channel of bandwidth KB. To reiterate, the number of pulses per second that can be transmitted over a channel is directly proportional to its bandwidth B.

The signal power S plays a dual role in information transmission. First, S is related to the quality of transmission. Increasing S, the signal power, reduces the effect of channel noise, and the information is received more accurately, or with less uncertainty. A larger signal-to-noise ratio (SNR) also allows transmission over a longer distance. In any event, a certain minimum SNR is necessary for communication. The second role of the signal power is not as obvious, although it is very important. We shall demonstrate that the channel bandwidth B and the signal power S are exchangeable; that is, to maintain a given rate and accuracy of information transmission, we can trade S for B, and vice versa. Thus, one may reduce B if one is willing to increase S, or one may reduce S if one is willing to increase B. In conclusion, the two primary communication resources are the bandwidth and the transmitted power.

Since the SNR is proportional to the power S, we can say that SNR and bandwidth are exchangeable.

The limitation imposed on communication by the channel bandwidth and the SNR is dramatically highlighted by Shannon's equation:

$$C = B \log_2 (1 + SNR) \text{ bit/s.} \quad (1.1)$$

Here C is the rate of information transmission per second. This rate C (known as the channel capacity) is the maximum number of binary symbols (bits) that can be transmitted per second with a probability of error arbitrarily close to zero. In otherwords, a channel can transmit $B \log_2 (1 + SNR)$ binary digits, or symbols, per second as accurately as one desires. Moreover, it is impossible to transmit at a rate higher than this without incurring errors. If there were no noise on the channel $(N = 0)$, $C \to \infty$, and communication would cease to be a problem. We could transmit any amount of information in the world over a channel.

It will be shown that the relationship between the bandwidth expansion factor and the SNR is exponential. Thus, if a given rate of information transmission requires a channel bandwidth B_1 and an SNR_1, then it is possible to transmit the same information over a channel bandwidth B_2 and a SNR_2, where

$$SNR_2 \simeq SNR_1^{B_1/B_2} \quad . \quad (1.2)$$

Thus, if we double the channel bandwidth, the required SNR is just a square root of the former SNR. In practice, the exchange between B and SNR is usually in the sense of increasing B to reduce transmitted power, and rarely the other way around.

Generally speaking, the transmission of signals in digital form comes much closer to the realization of the limit (1.2) than does the transmission of signals in analog form.

4. Modulation

Baseband signals produced by various information sources are not always suitable for direct transmission over a given channel. These signals are usually further modified to facilitate transmission. This conversion process is known as modulation. In this process, the baseband signal is used to modify some parameter of a high-frequency carrier signal.

A carrier is a sinusoid of high frequency, and one of its parameters - such as amplitude, frequency, or phase - is varied in proportion to the baseband signal $m(t)$.

At the receiver, the modulated signal must pass through a reverse process called demodulation in order to reconstruct the baseband signal.

As mentioned earlier, modulation is used to facilitate transmission. Some of the important reasons for modulations are:

- Ease of radiation
- Simultaneous transmission of several signals
- Effecting the exchange of SNR with B