Lecture(0): Information Theory

Abdelrahman Elnamaki

September 16, 2024

Introduction

In the field of information theory, understanding the flow of information through a communication system is crucial. A typical communication system comprises several key components that work together to ensure the accurate transmission of information. These components include the source, encoder, transmission medium, channel, noise, decoder, and receiver. Each component serves a specific function:

1 System Components

1. **Source (S):** The source is the origin of the information to be communicated. It could represent various types of data, such as text, voice, or sensor readings. The role of the source is to generate the message, M, that is to be transmitted to the receiver.

2. **Encoder:** The encoder is responsible for converting the message from the source into a format that is suitable for transmission. This process may involve data compression, binary encoding, or other forms of signal processing to ensure that the message can be efficiently and accurately sent through the transmission medium. The encoded message is typically represented as $E(M)$.

3. **Transmission:** Once the message is encoded, it is transmitted through a physical or wireless medium. This step represents the propagation of the signal from the source toward the receiver. The transmission medium could include fiber optic cables, radio waves, or other forms of communication channels.

4. **Channel:** The channel is the path through which the encoded message travels. It represents the physical infrastructure that carries the signal between the source and the receiver. However, real-world communication channels are not perfect and often introduce interference.

5. **Noise (N):** Noise refers to any form of disturbance or interference that affects the message as it travels through the channel. This could include electrical noise in cables, atmospheric interference in wireless communication, or packet loss in data networks. The presence of noise can result in the distortion of the message, M', such that the received message may differ from the original.

6. **Received Message (R):** The received message is the signal that reaches the receiver after passing through the channel. Due to the presence of noise, R may not be identical to the original encoded message, and errors may occur. The receiver's goal is to interpret the message despite these imperfections.

7. **Decoder:** The decoder processes the received message, R, to reconstruct the original message, M. This step involves reversing the encoding process and correcting for any errors introduced by noise in the channel. A well-designed decoder aims to recover the message with minimal loss of information.

8. **Receiver (S'):** The receiver is the final component of the system and is responsible for interpreting the decoded message. The objective is for the receiver to reconstruct M as accurately as possible. In an ideal communication system, the message received by the receiver, M', would be identical to the message sent by the source, M. However, due to noise, this is not always possible.

2   Bit Flip Probability Model

In a bit-flipping model:

- The probability of a bit **staying 0** (given that its input was 0) is $1 - f$,

- The probability of a bit **flipping to 1** (given that its input was 0) is f,

- The probability of a bit **staying 1** (given that its input was 1) is $1 - f$,

- The probability of a bit **flipping to 0** (given that its input was 1) is f.

This describes a *noisy channel model*, where each bit has a probability f of flipping, regardless of its initial value, and a probability $1 - f$ of staying the same.

2.1   Formalizing the Model

Let:

- X be the original bit (either 0 or 1),

- Y be the received or flipped bit,

- f be the flip probability.

Then, the conditional probabilities are:

- If $X = 0$:

$$P(Y = 0|X = 0) = 1 - f,$$
$$P(Y = 1|X = 0) = f.$$

- If $X = 1$:

$$P(Y = 1|X = 1) = 1 - f,$$
$$P(Y = 0|X = 1) = f.$$

2.2 Transition Matrix

A convenient way to represent this kind of bit-flip model is through a *transition matrix*. If X is the input and Y is the output, the matrix can be written as:

$$\begin{bmatrix} P(Y = 0|X = 0) & P(Y = 1|X = 0) \\ P(Y = 0|X = 1) & P(Y = 1|X = 1) \end{bmatrix} = \begin{bmatrix} 1 - f & f \\ f & 1 - f \end{bmatrix}$$

This matrix describes how likely it is for each bit to either flip or stay the same based on the flip probability f.

2.3 Summary

- The probability of flipping a bit (from 0 to 1 or from 1 to 0) is f,

- The probability of the bit remaining the same is $1 - f$.

3 Redundancy in Communication Systems

Redundancy refers to the inclusion of additional information in a message to improve its reliability and accuracy, particularly in the presence of noise. It helps detect or correct errors during transmission and ensures that the message can still be interpreted correctly.

3.1 Types of Redundancy

- **Error Detection Codes:** These include parity bits, checksums, and cyclic redundancy checks (CRC). The parity bit is a simple example where a bit is added to ensure the number of 1s is even (even parity) or odd (odd parity). For a message M with n bits, the parity bit P is calculated as:

$$P = \text{parity}(M) = \left(\sum_{i=1}^{n} M_i \right) \mod 2$$

3

- **Error Correction Codes (ECC):** These codes allow both detection and correction of errors. For example, a Hamming code uses parity checks for multiple subsets of bits. For a given data bit d_i in an $[n, k]$ code, where n is the total number of bits and k is the number of data bits, the Hamming distance d_H determines the error correction capability:

$$d_H = 2t + 1$$

where t is the number of correctable errors.

- **Repetition Codes:** This simple form of redundancy involves repeating each bit multiple times to ensure accuracy. For instance, in a repetition code with r repetitions, the probability of error decreases with the number of repetitions. The probability of error P_e in a repetition code is given by:

$$P_e = \binom{r}{\lfloor r/2 \rfloor} p^{\lfloor r/2 \rfloor} (1 - p)^{\lceil r/2 \rceil}$$

where p is the probability of a bit error.

- **Forward Error Correction (FEC):** FEC codes add redundancy so that the receiver can correct errors without needing retransmissions. An example is the Reed-Solomon code, which can correct up to t symbol errors:

$$t = \frac{n - k}{2}$$

where n is the code length and k is the number of data symbols.

- **Interleaving:** This technique rearranges data to minimize the impact of burst errors. For a block of data, interleaving rearranges the bits or symbols to spread out errors. The interleaved sequence I can be expressed as:

$$I(i) = \text{data}\left(\left\lfloor \frac{i}{L} \right\rfloor + (i \mod L) \cdot k \right)$$

where L is the length of the interleaving block and k is the number of interleaving rows.

- **Hybrid ARQ:** Combines error detection with retransmission requests and FEC. The system uses error detection to request retransmission when errors are detected, and FEC to correct errors in the received data. The efficiency of Hybrid ARQ can be represented by the effective code rate R_{eff}:

$$R_{eff} = \frac{C}{C + T}$$

where C is the number of correctable bits and T is the total transmitted bits including redundancy.

Redundancy increases the reliability of communication systems but must be balanced with efficiency to avoid excessive bandwidth usage or processing delays.

4 Channel Capacity

Channel capacity is a measure of the maximum rate at which information can be transmitted over a communication channel with a specified level of noise, ensuring that the transmission is as accurate as possible. It represents the highest data rate that can be achieved with an arbitrarily low probability of error, given optimal encoding and decoding schemes.

4.1 Shannon's Capacity Theorem

Shannon's capacity theorem provides a formula to calculate the channel capacity C for a continuous channel with Gaussian noise. The formula is:

$$C = B \log_2 \left(1 + \frac{S}{N} \right)$$

where:

- C is the channel capacity in bits per second (bps).

- B is the bandwidth of the channel in hertz (Hz).

- S is the average signal power in watts (W).

- N is the average noise power in watts (W).

4.2 Explanation

- **Bandwidth (B):** The range of frequencies that the channel can transmit. A larger bandwidth allows for higher data rates.

- **Signal-to-Noise Ratio (S/N):** The ratio of the signal power to the noise power. A higher signal-to-noise ratio means better transmission quality and a higher channel capacity.

- **Logarithm Function:** The capacity increases logarithmically with the signal-to-noise ratio. This reflects that gains in capacity become smaller as the signal-to-noise ratio increases.

The Shannon-Hartley theorem is derived from information theory concepts, including:

- **Information Entropy:** Measures the average amount of uncertainty or information content. For continuous signals, entropy is:

$$H(X) = - \int p(x) \log_2 p(x) \, dx$$

where $p(x)$ is the probability density function of the signal.

- **Mutual Information:** Represents the amount of information obtained about one variable through another. For a channel, it is:

$$I(X;Y) = H(Y) - H(Y|X)$$

where $H(Y)$ is the entropy of the output and $H(Y|X)$ is the conditional entropy of the output given the input.

In summary, channel capacity C is the maximum rate at which information can be reliably transmitted over a channel, and it depends on both the channel's bandwidth and the signal-to-noise ratio.

5   Capacity of a Binary Symmetric Channel (BSC)

A Binary Symmetric Channel (BSC) is a communication channel where the input and output are binary (0 or 1) and errors occur with a certain probability. Specifically, the channel flips a transmitted bit with probability p, and transmits it correctly with probability $1 - p$.

5.1   Capacity of the Binary Symmetric Channel

The capacity C of a Binary Symmetric Channel is the maximum rate at which information can be transmitted over the channel with an arbitrarily low probability of error. The capacity is given by:

$$C = 1 - H(p)$$

where $H(p)$ is the binary entropy function defined as:

$$H(p) = -p \log_2 p - (1 - p) \log_2 (1 - p)$$

5.2   Explanation

- **Binary Entropy Function $H(p)$:** Measures the uncertainty or randomness associated with the probability p. It reaches its maximum when $p = 0.5$, representing the highest uncertainty (i.e., the channel is most noisy).

- **Channel Capacity C:** Reflects the maximum rate of reliable transmission. It decreases as the error probability p increases. When $p = 0$ or $p = 1$, the capacity is 1 bit per channel use (no errors), while for $p = 0.5$, the capacity is 0 (maximum uncertainty).

The formula captures how the capacity of the channel is impacted by the probability of errors, providing a measure of the channel's efficiency in transmitting information reliably.

6   Conclusion

Understanding the components and models of communication systems helps optimize the transmission of information. By employing redundancy and error correction techniques, we can ensure accurate and efficient communication despite the presence of noise. The principles of entropy and channel capacity are fundamental in designing and analyzing communication systems, providing a framework for achieving reliable data transmission.