ELC5693-GENERAL COMMUNICATION THEORY
6B07112 Electronic and Electrical Engineering

СӘТБАЕВ УНИВЕРСИТЕТІ SATBAYEV UNIVERSITY

Lecture 12: Fundamentals of Coding Theory

Dosbayev Zhandos Makhsutuly, senior lecturer
E-mail: zh.dosbayev@satbayev.university

Outline

Introduction to Coding Theory

Error Detection and Correction

Shannon's Theorem and Channel Capacity

Linear Codes and Generator Matrices

Hamming Codes and Parity-Check Matrices

Cyclic Codes and Polynomial Representation

Applications

Fundamentals of Coding Theory

Welcome to Lecture 12, where we delve into the fascinating world of Coding Theory. This fundamental branch of information theory and mathematics plays a crucial role in our modern digital landscape. In this lecture, we'll explore the core principles that underpin reliable data transmission and storage across various communication channels.

Coding theory is the backbone of our digital communication systems, ensuring that information can be transmitted accurately and efficiently, even in the presence of noise and errors. As we progress through this lecture, we'll uncover the elegant mathematical concepts and practical applications that make coding theory an indispensable tool in our increasingly connected world.

Introduction to Coding Theory

1

Origins

Coding theory emerged in the mid-20th century, pioneered by Claude Shannon's groundbreaking work in information theory. Shannon's 1948 paper "A Mathematical Theory of Communication" laid the foundation for modern digital communication and data compression.

2

Core Objectives

The primary goals of coding theory are to devise efficient methods for encoding information, detecting errors in transmitted data, and correcting these errors to ensure reliable communication across noisy channels.

3

Interdisciplinary Nature

Coding theory draws from various mathematical disciplines, including algebra, combinatorics, and probability theory. It finds applications in diverse fields such as telecommunications, data storage, and cryptography.

Importance of Coding Theory

Reliable Communication

Coding theory ensures that information can be transmitted accurately over noisy channels, minimising data corruption and loss. This is crucial for everything from mobile phone calls to deep space communications.

Data Integrity

In data storage systems, coding theory helps maintain the integrity of stored information, protecting against bit rot and hardware failures. This is essential for long-term data preservation and reliability.

Efficient Transmission

By employing data compression techniques, coding theory enables more efficient use of bandwidth, allowing for faster data transmission and reduced storage requirements.

Error Correction

Coding theory provides mechanisms for detecting and correcting errors in transmitted or stored data, ensuring that information remains accurate even in challenging environments.

Fundamental Concepts: Codewords, Codes, and Encoding

Codewords

Codewords are the fundamental units in coding theory. They are sequences of symbols from a defined alphabet (e.g., binary digits) that represent encoded information. Each valid codeword belongs to a specific code and carries a unique message.

Codes

A code is a set of codewords that adhere to specific rules or properties. Codes can be categorised based on their structure (e.g., linear codes, cyclic codes) and their error-correcting capabilities. The design of efficient codes is a central challenge in coding theory.

Encoding

Encoding is the process of transforming raw information into codewords according to a predefined scheme. This process often involves adding redundancy to the original data, which later aids in error detection and correction during decoding.

Error Detection and Correction

1

Error Detection

Error detection techniques identify when transmitted data has been corrupted. Common methods include parity checks and cyclic redundancy checks (CRC). These techniques add extra bits to the data, allowing the receiver to verify its integrity.

2

Error Correction

Error correction goes a step further by not only detecting errors but also reconstructing the original data. Techniques like Hamming codes and Reed-Solomon codes can correct a certain number of bit errors without requiring retransmission.

3

Forward Error Correction

FEC techniques preemptively add redundancy to data before transmission, allowing the receiver to correct errors without requesting retransmission. This is crucial for applications where real-time communication is essential or retransmission is impractical.

Shannon's Theorem and Channel Capacity

1

Shannon's Noisy-Channel Coding Theorem

This fundamental theorem, proposed by Claude Shannon in 1948, establishes the maximum rate at which information can be transmitted reliably over a noisy communication channel. It provides a theoretical upper bound on the performance of error-correcting codes.

2

Channel Capacity

Channel capacity is the tightest upper bound on the amount of information that can be reliably transmitted over a communications channel. It's measured in bits per channel use and depends on the channel's bandwidth and signal-to-noise ratio.

3

Implications for Coding Theory

Shannon's theorem suggests that it's possible to achieve error-free communication up to the channel capacity using appropriate coding techniques. This has driven the development of increasingly sophisticated error-correcting codes.

4

Practical Considerations

While Shannon's theorem provides a theoretical limit, practical systems often operate below this capacity due to complexity constraints and the need for reasonable encoding/decoding times.

Linear Codes and Generator Matrices

Property	Linear Codes	Generator Matrices
Definition	Codes where any linear combination of codewords is also a codeword	Matrices used to encode messages into codewords for linear codes
Key Feature	Closed under addition and scalar multiplication	Rows form a basis for the code space
Advantage	Efficient encoding and decoding algorithms	Compact representation of the entire code
Example	Hamming codes, Reed-Solomon codes	G = [I | P], where I is identity matrix and P is parity matrix

Hamming Codes and Parity-Check Matrices

Hamming Codes

Hamming codes are a family of linear error-correcting codes named after Richard Hamming. They are capable of detecting up to two simultaneous bit errors and correcting single-bit errors. The most common Hamming code is (7,4), which encodes 4 data bits into 7 bits.

Parity-Check Matrices

Parity-check matrices are fundamental to the structure of linear codes. For a code C, the parity-check matrix H is defined such that for any codeword c in C, H * c^T = 0. This property is used for both encoding and decoding, particularly in syndrome decoding.

Relationship

For Hamming codes, the parity-check matrix is designed so that each column is unique and non-zero. This structure allows for efficient error detection and correction by identifying the position of a flipped bit through syndrome calculation.

Cyclic Codes and Polynomial Representation

1

Cyclic Codes

Cyclic codes are a subclass of linear codes with the property that any cyclic shift of a codeword is also a codeword. This structure allows for efficient encoding and decoding using shift registers.

2

Polynomial Representation

Cyclic codes can be elegantly represented using polynomials over finite fields. Each codeword is associated with a polynomial, and the code itself is defined by a generator polynomial g(x).

3

Encoding Process

To encode a message m(x), it is multiplied by x^(n-k) and then divided by g(x). The remainder of this division is added to x^(n-k)m(x) to form the codeword.

4

Decoding and Error Correction

Decoding cyclic codes often involves syndrome calculation and error-locator polynomials. Techniques like the Berlekamp-Massey algorithm are used for efficient decoding.

Applications of Coding Theory

Telecommunications

Coding theory is crucial in satellite communications, mobile networks, and deep space communications, ensuring reliable data transmission over vast distances and through noisy environments.

Data Storage

Error-correcting codes are used in hard drives, SSDs, and optical discs to maintain data integrity over long periods and protect against physical degradation of storage media.

Wireless Networks

Wi-Fi, Bluetooth, and other wireless protocols rely heavily on coding theory to ensure reliable data transmission in the presence of interference and signal attenuation.

Bioinformatics

Coding theory techniques are applied in DNA sequencing and analysis, helping to correct errors in genetic data and improve the accuracy of genomic studies.