S-72.3410 Coding Methods (5 cr) P

Lectures: Mondays 9–12, room E110, and Wednesdays 9–12, hall S4 (on January 30th this lecture will be held in E111!)

Teacher: Markku Liinaharja, room 217, Otakaari 7A, tel. 451 2172, e-mail: Markku.Liinaharja@tkk.fi

Tutorials: Tuesdays 10–12, room E110, and Fridays 12–14, hall S2, first tutorial on Friday January 25th.

Assistant: Esa Seuranen, room I446 (Otakaari 5; entrance via I447), tel. 451 5861, e-mail: eseurane@cc.hut.fi

Home page:
`http://www.comlab.hut.fi/studies/3410/index.html`

Contents

▷ Coding theory
▷ Coding and decoding algorithms
▷ Application to digital communication and storage

Cryptography is *not* considered in this course. The interested students are referred to T-79.4501 Cryptography and Data Security *or* T-79.5501 Cryptology.

Prerequisites

Recommended: S-72.2410 Information Theory.

More importantly: A good mathematical background (algebra, linear algebra) *or* an interest in mathematics.

Literature (1)

The following books are particularly appropriate for students in electrical engineering:

[Wic] S. B. Wicker, *Error Control Systems for Digital Communication and Storage*, Prentice-Hall, Upper Saddle River, NJ, 1995. [**Course literature**.]

J. Castiñeira Moreira & P. G. Farrell, *Essentials of Error-Control Coding*, Wiley, Chichester, UK, 2006. [**Course literature** for turbo and LDPC codes.]

S. Lin and D. J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*, Second edition, Pearson Prentice Hall, Upper Saddle River, NJ, 2004.

Literature (2)

E. R. Berlekamp, Algebraic Coding Theory, rev. ed., Aegean Park Press, Laguna Hills, 1984.

F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*, North-Holland, Amsterdam, 1977. [The "bible" of coding theory.]

V. Pless, *Introduction to the Theory of Error-Correcting Codes*, Wiley, New York, 1989.

J. H. van Lint, *Introduction to Coding Theory*, 3rd ed., Springer-Verlag, New York, 1999.

Journals and Conferences

Research in coding theory is carried out in both the information theory and the discrete mathematics society.

Main journals: *IEEE Transactions on Information Theory*; and *Designs, Codes and Cryptography.*

Other journals: *Discrete Mathematics*, etc.

Main conference: *IEEE Symposium on Information Theory.*

Outline of the Course

1. Introduction (1)
2. Fields and polynomials over fields (2)
3. Linear block codes, cyclic codes (3)
4. BCH and Reed-Solomon codes (1)
5. Convolutional codes, the Viterbi algorithm (2)
6. Channels with feedback (1)
7. Turbo codes (1) (Prof. Östergård)
8. LDPC codes (1) (Prof. Östergård)

To pass the course: Pass the exam and solve at least 30 % of the homework problems.

Digital Communication Systems (1)

A digital communication system is a means of transporting information from one party (A) to another (B).

digital The system uses a sequence of symbols from a finite alphabet ($V_q = \{0, 1, \ldots, q-1\}$) to represent the information. Transmission in digital form allows for error control coding.

The basic elements of a digital communication system are depicted in [Wic, Fig. 1-1].

Digital Communication Systems (2)

The modulator maps the information symbols onto signals that can be efficiently transmitted over the communication channel. The selection of a modulation format is a complex subject that is outside the scope of this course; this subject is treated in various other S-72 courses.

The physical channel attenuates the transmitted signal and introduces noise. The most commonly assumed noise model is the *additive white Gaussian noise* (AWGN) model.

In environments where noise is present, the demodulated data contains errors. This is usually characterized in terms of a *bit error rate* (BER).

Design Criteria for Communication Systems

- BER
- Throughput
- Complexity
- Cost
- Weight
- Heat dissipation
- Fault tolerance

Shannon and Information Theory

In the late 1940s, the work of Shannon and Hamming at Bell Laboratories laid the foundation for error control coding.

Shannon The father of *information theory* with his paper “A Mathematical Theory of Communication” in 1948. Proved the limits for ideal error control.

Hamming Presented and analyzed the first practical error control system (based on Hamming codes).

Shannon’s results in his seminal paper allows to determine the minimum possible number of symbols necessary for the error-free representation of a given message. A longer message containing the same information is said to have *redundant* symbols.

Code Types

source codes Remove *uncontrolled redundancy* and format data.

secrecy codes Encrypt information so that the information cannot be understood by anyone except the intended recipient(s).

error control codes (or channel codes) Format the transmitted information so as to increase its immunity to noise by inserting *controlled redundancy* into the information stream.

Integration of these codes into the basic communication system model is depicted in [Wic, Fig. 1-3].

▷ The order of the three codes is crucial!

Strategies with Error Control Codes

When error control codes are used, there are several possible ways of reacting to a detected error:

1. Request a retransmission of the erroneous word.
2. Tag the word as being incorrect and pass it along.
3. Attempt to correct the errors in the received word.

The Noisy Channel Coding Theorem

Theorem 1-1. With every channel we can associate a "channel capacity" C. There exist error control codes such that information can be transmitted across the channel at rates less than C with arbitrarily low bit error rate.

▷ The proof of this theorem is existential!

A Parity-Check Code (1)

Take the output of a binary source and break it up into k-bit blocks of the form

$$\mathbf{m} = (m_0, m_1, \ldots, m_{k-1}).$$

At the end of every such block, append a redundant bit b as follows to get a codeword (in the rest of this lecture, all additions are carried out modulo 2):

$$\mathbf{c} = (m_0, m_1, \ldots, m_{k-1}, b), \text{ where } b = \sum_{i=0}^{k-1} m_i.$$

A Parity-Check Code (2)

The receiver adds together the values in each coordinate. If the sum is 1, we know that the received word is in error.

⇒ All erroneous words that contain an odd number of errors are detected.

This code is a *single-error-detecting* (or 1-error-detecting) code. A code is said to be *t-error-detecting* if *all* erroneous words with at most t errors are detected.

A Hamming Code (1)

Let the message blocks and codeword be as follows (the length of these is 4 and 7, respectively):

$$\begin{aligned}
\mathbf{m} &= (m_0, m_1, m_2, m_3),\\
\mathbf{c} &= (m_0, m_1, m_2, m_3, b_0, b_1, b_2),\\
b_0 &= m_1 + m_2 + m_3,\\
b_1 &= m_0 + m_1 + m_3,\\
b_2 &= m_0 + m_2 + m_3.
\end{aligned}$$

A Hamming Code (2)

The received word is denoted by **r**. The following (binary) values are computed for each such word:

$$\begin{aligned}
\mathbf{r} &= (r_0, r_1, r_2, r_3, r_4, r_5, r_6),\\
s_0 &= r_1 + r_2 + r_3 + r_4,\\
s_1 &= r_0 + r_1 + r_3 + r_5,\\
s_2 &= r_0 + r_2 + r_3 + r_6.
\end{aligned}$$

A Hamming Code (3)

If $s_0 = s_1 = s_2 = 0$, then the received word is a valid word. Otherwise, the value of (s_0, s_1, s_2) gives the position of a single error:

(s_0, s_1, s_2)	Error location
000	None
001	r_6
010	r_5
011	r_0
100	r_4
101	r_2
110	r_1
111	r_3

A Hamming Code (4)

This code is capable of correcting all received words with at most 1 error ⇒ it is a *single-error-correcting* code. Analogously, a code is said to be *t-error-correcting* if all erroneous words with at most t errors can be corrected.

Performance Improving with Error Control (1)

Error control is achieved via redundancy. The *code rate* R denotes the ratio of k, the number of data symbols transmitted per code word, to n, the number of symbols transmitted per code word. The code in the previous example has rate 4/7.

If we want the symbol rate R_S to remain constant, the overall transmission rate must be increased to R_S/R.

⇒ If the transmission power level is constant, then the received energy per symbol is reduced from E_S to RE_S.

Performance Improving with Error Control (2)

The demodulated BER is then increased with respect to its previous value! However, if the code is well selected, the BER at the output of the decoder is better than with the original, uncoded system.

coding gain The additional transmitted power that is required to obtain the same performance without coding.

$$P_{\mathrm{dB}} = 10 \log_{10} P_{\mathrm{watts}}$$

Other Applications of Codes

In this course, codes for error-detecting and error-correcting purposes are discussed. In general, a code is any subset of words in a discrete space, and there is a wide variety of possible applications.

Example. A binary *covering* code: $C = \{0000, 0101, 1110, 1011\}$. For any binary word x of length 4, there exists a word in C that differs from x in at most one coordinate. Applications:

- (Lossy) data compression
- Systems for betting (football pools)