ECEN 604 - Channel Coding for Communication
Lectures by Graduate Students

Narayanan Rengaswamy
Department of Electrical and Computer Engineering
Texas A&M University, College Station, TX 77843

October 2015

Lecture 1

1 Introduction

A generic communication system block diagram is given in Fig. 1. A source generates digital data, which is passed through a source encoder that compresses it so that it can be represented with minimum number of bits/symbols. Then, the channel encoder adds calculated redundancy to the data in order to counter the noise introduced by the channel. Subsequently, the codeword is transmitted over the noisy channel W. At the receiver, the channel decoder receives a noisy observation of the codeword and attempts to decode it given the encoding scheme employed by the channel encoder at the transmitter. The average number of times the output of the decoder has at least one bit/symbol in error is represented as the block error rate P_B. Then, the source decoder tries to reconstruct the actual data and sends it to the sink.

In this course, we are interested in the channel coding blocks: channel encoder and decoder. We will be addressing the following question:

How do we design efficient codes for reliable transmission over a noisy channel at a given rate, with minimum latency and complexity in the system?

1.1 Trade-offs in Communication Design

There are four major parameters that are of interest for the design of a communication system:

- Rate: the amount of information transmitted per use of the channel.
- Reliability: the probability of decoding error, often termed the block/word error rate.

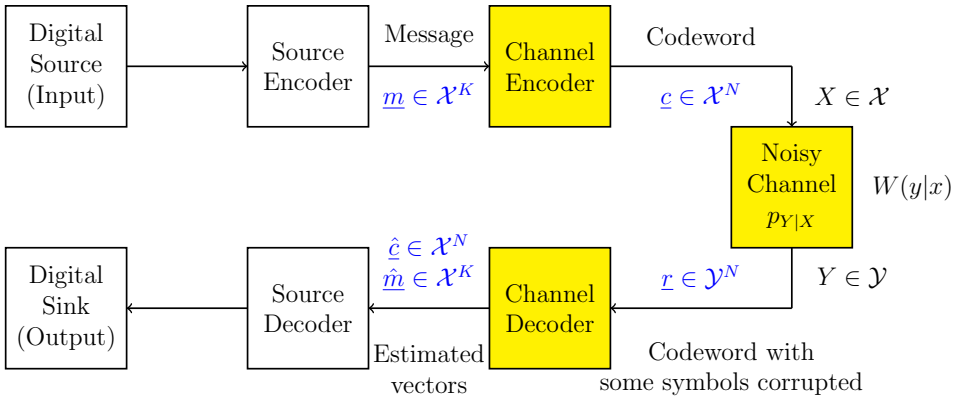


Figure 1: A communication system highlighting the main blocks from a coding theoretic perspective. It is slightly more detailed than the model considered by Shannon in [1] in the sense that the source and channel coding blocks are shown for emphasis.

- Complexity: the number of computations involved in the encoding and decoding processes.
- Latency: the time it takes to decode per transmission.

For now, let us ignore complexity and latency. Then, we have a trade-off between rate and reliability. If the channel is close to ideal, i.e. very reliable, then we can transmit more messages and still be able to receive them correctly in a given period of time. However, if the channel is very noisy, i.e. less reliable, then with high probability the receiver will not be able to decipher the message that was transmitted. Therefore, we will have to transmit the same message multiple times until we are successful. Just imagine repeating the same thing in a conversation if the surrounding is very noisy. Clearly, just repetition is a very inefficient method of utilizing the communication resources.

Channel coding helps in reducing re-transmissions and increasing reliability over a noisy channel. The way it is done is by adding calculated redundancy to the data so that even if some bits/symbols are corrupted, we will still be able to recover the information by leveraging the redundancy. Now, let us see a simple channel model to understand some preliminary channel coding methods.

1.2 Binary Symmetric Channel

The binary symmetric channel is defined by one parameter, p, which is the probability that a transmitted bit gets flipped. The channel is denoted as BSC(p) and is represented as given in Fig. 2. In terms of the channel transition probabilities, the BSC is defined as

$$W(y|x) \triangleq P(Y = y|X = x) = \begin{cases} 1-p & , \text{ if } y = x \\ p & , \text{ if } y \neq x \end{cases} \quad \forall\, x \in \mathcal{X}, y \in \mathcal{Y}. \tag{1}$$

This is the one of most widely used channel models for a variety of applications. The generality of the model is (hopefully) evident.

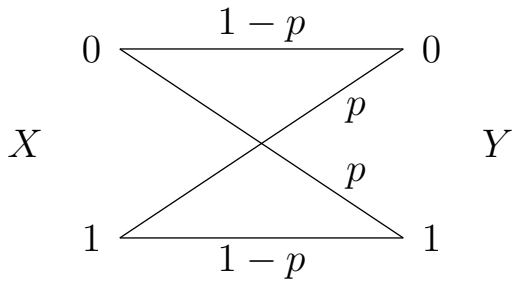


Figure 2: Binary Symmetric Channel with crossover probability p. X and Y are random variables that represent the transmitted and received bits, respectively.

1.3 Comparison of Simple Transmission Schemes

On the BSC(p), let us compare three simple schemes for transmitting 6 bits using 6 independent uses of the channel, in terms of their probability of decoding error P_{B}.

Uncoded Transmission:

The rate of transmission is 1. The received information is correct iff all the bits are not flipped. Hence, the probability of error can be calculated as

$$P_{\mathrm{B}} = 1 - (1-p)^6 \approx 6p = O(p).$$

We see that even though the rate is maximum, the reliability is worse.

$(3, 1)$ Repetition Code:

Every information bit is transmitted thrice. So, effectively, the rate of (coded) transmission is 1/3. The decoder will wait for every 3 bits and decode the information as that bit which appears in majority out of the 3 bits. Clearly, for an information bit to be wrongly decoded, at least two of the three transmitted bits must be flipped. Hence, the probability of decoding error can be calculated as

$$\begin{aligned} P_{\mathrm{B}} &= 1 - (1 - \text{Prob.[error in either of the 2 decoded message bits]})^2 \\ &= 1 - \left(1 - \left(\binom{3}{2} p^2(1-p) + p^3\right)\right)^2 = O(p^2). \end{aligned}$$

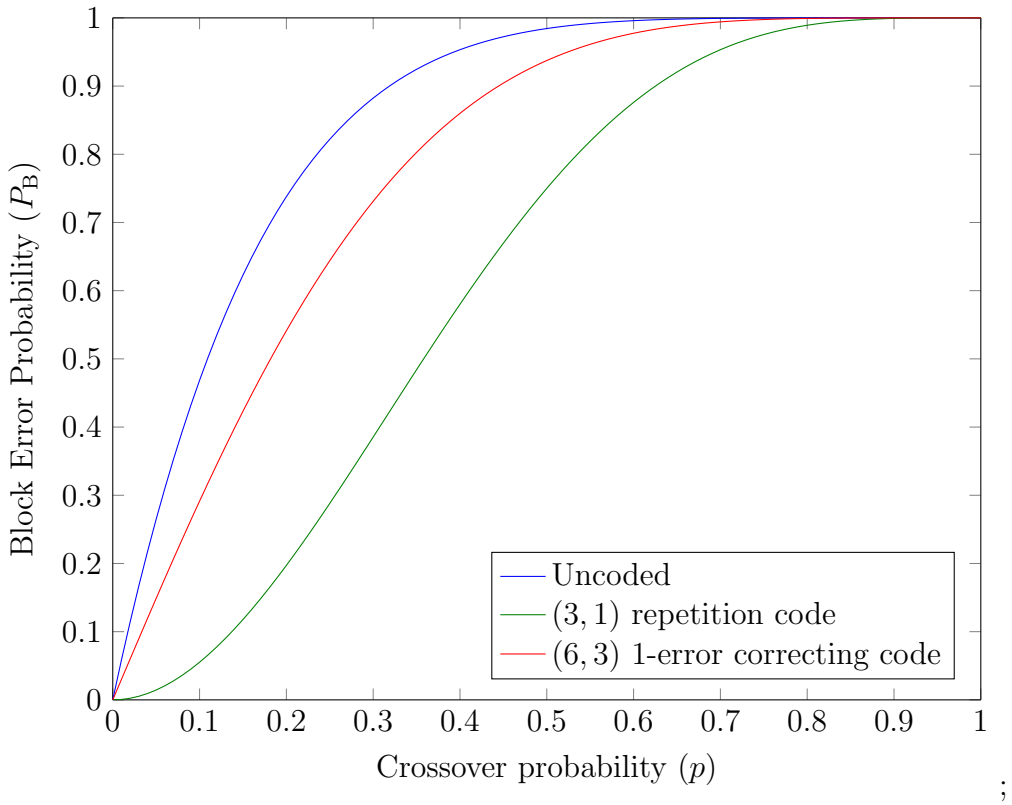
This is the other extreme where the rate is low but the reliability is much better.

A $(6, 3)$ 1-Error Correcting Code:

This is a code which has $2^3 = 8$ codewords in the code book such that any two codewords differ in at least 3 bits. You can try to list such a set of eight vectors. We will see later that this 3 is the minimum distance, $d_{\min}$, of the code. It can be shown that this code can correct at most one error in the received vector. Therefore, the probability of decoding error is given as

$$P_{\mathrm{B}} = 1 - (1-p)^6 - \binom{6}{1} p(1-p)^5.$$

Interestingly, we now have a code that has an acceptable rate, 0.5, with a reliability in between the above two schemes. The three schemes are compared in Fig. 3. Evidently, from this example, we see the trade-off between rate and reliability.



;

Figure 3: A comparison of three coding schemes on the BSC. The uncoded scheme has rate 1, the repetition code has rate 1/3 and the $(6, 3)$ code has rate 1/2. The trade-off between rate and reliability is evident.

Now, the question is how do we design codes for a given application that requires a certain amount of reliability with a tolerance on the rate?

This is the motivation to build mathematical tools that enable us to construct codes for a target application, given the requirements, and this art is called *channel coding*. Shannon's information theory precisely gives a definition for *channel capacity*, which is the maximum rate at which one can communicate over a channel at an arbitrarily low reliability. For example, the capacity of BSC(p) is given by

$$C_{\text{BSC}} = 1 - h_2(p) = 1 - p\,\log_2\left(\frac{1}{p}\right) - (1-p)\,\log_2\left(\frac{1}{1-p}\right), \tag{2}$$

where $h_2(\cdot)$ is the binary entropy function. Throughout this course, logarithms will be base-2 unless specified otherwise. In his seminal paper in 1948, titled *A Mathematical Theory of Communication* [1], Shannon quantified information, defined the notion of capacity, and proved that capacity-achieving codes exist through his brilliant random coding arguments. But, in practice, we need deterministic code constructions that operate close to capacity. Designing such codes has been the pursuit of coding theorists for over 60 years.

1.4 Credit Card Number Example

To get a feel for coding, let us look at the coding scheme for credit card numbers. Digit sum of a non-negative integer is given by

$$d_{\text{sum}}(x) = x \bmod 9, \text{ unless } x = 9. \tag{3}$$

Let $x_1, x_2, \ldots, x_{16}$ denote the digits of the credit card number, starting from the first digit. Then, the check for the credit card number is

$$\left[\sum_{i=1}^{8} x_{2i} + \sum_{i=1}^{8} d_{\text{sum}}(2x_{2i-1})\right] \bmod 10 = 0. \tag{4}$$

It can be shown that this code corrects all single digit errors and all adjacent transpositions.

1.5 Important Notations and Definitions

$\mathcal{X}$	Input alphabet of the channel
$\mathcal{Y}$	Output alphabet of the channel
K	Number of message symbols per transmission block (or) Dimension of the code
N	Number of code symbols per transmission block (or) Length of the code
R	Rate of the code
M	Size of the code, i.e. number of codewords in the code book
$\underline{m} \in \mathcal{X}^K$	Message vector
$\underline{c} \in \mathcal{X}^N$	Codeword vector
$\underline{e} \in \mathcal{Y}^N$	Error vector
$\underline{r} \in \mathcal{Y}^N$	Received vector; a noisy observation of $\underline{c}$

$\underline{\hat{c}} \in \mathcal{X}^N$	Estimated codeword vector
$\underline{\hat{m}} \in \mathcal{X}^K$	Estimated message vector
$W(y\|x)$	Channel transition probability from input $x \in \mathcal{X}$ to output $y \in \mathcal{Y}$
BSC(p)	Binary Symmetric Channel with crossover probability p
BEC(ϵ)	Binary Erasure Channel with erasure probability ϵ
BIAWGNC(σ)	Binary Input Additive White Gaussian Noise Channel with noise variance σ^2
P_b	Bit Error Probability
P_B	Block Error Probability
$\mathbb{F}_q$	Finite field $\mathbb{F}$ with q elements
$\mathbb{R}$	The set of all real numbers
$\mathbb{Z}$	The set of all integers

Some important definitions are given below:

1. Rate of a code, $R = \dfrac{K}{N}$.
2. Total number of codewords in a code, $M \geq |\mathcal{X}|^K$.
3. Received vector, $\underline{r} = \underline{c} + \underline{e}$. The addition in $\mathbb{F}_q$ is $(\mathrm{mod} q)$ if q is prime.
4. Memoryless channel: $W(y_1^N|x_1^N) = \prod_{i=1}^{N} W(y|x)$ for $W : \mathcal{X} \to \mathcal{Y}$, i.e. every use of the channel is independent of all the previous uses.
5. Binary Symmetric Channel, BSC(p), $W : \{0,1\} \to \{0,1\}$ is defined as

$$W(y|x) \triangleq P(Y = y|X = x) = \begin{cases} 1-p & , \text{ if } y = x \\ p & , \text{ if } y \neq x \end{cases} \quad \forall\, x \in \mathcal{X}, y \in \mathcal{Y}. \tag{5}$$

6. Binary Erasure Channel, BEC(ϵ), $W : \{0,1\} \to \{0,1,?\}$ is defined as

$$W(y|x) \triangleq P(Y = y|X = x) = \begin{cases} 1-\epsilon & , \text{ if } y = x \\ \epsilon & , \text{ if } y =? \end{cases} \quad \forall\, x \in \mathcal{X}, y \in \mathcal{Y}. \tag{6}$$

7. Binary input AWGN Channel, BIAWGNC(σ), $W : \{-1,1\} \to \mathbb{R}$ is defined as follows:

$$Y = X + E, \text{ where } E \sim \mathcal{N}(0, \sigma^2).$$

   Here, E is a random variable that represents the noise and is zero-mean Gaussian with variance $\sigma^2 = N_0/2$. Using this, the channel transition probabilities are defined as

$$W(y|x) \triangleq P(Y = y|X = x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{|y-x|^2}{2\sigma^2}} \quad \forall\, x \in \mathcal{X}, y \in \mathcal{Y}. \tag{7}$$

1.6 Applications of Error-Correction Coding

Some applications for ECC are:

1. Cell phones.
2. Bluetooth, WiFi.
3. Space communications.
4. Global Positioning System (GPS).
5. Digital Video Broadcast (DVB).
6. Storage devices, e.g. CD, DVD, hard drive, flash memories.
7. Credit card numbers, ISBN numbers etc.

1.7 Overview of the Course

In this introductory course to channel coding, we will understand algebraic codes in good detail and look at their applications. We will begin with binary linear codes, then work our way towards non-binary cyclic codes over finite fields, which will lead us to the highly useful Reed-Solomon (RS) and Bose-Chaudhuri-Hocquenghem (BCH) codes. We will spend a considerable amount of time for RS codes and try to cover BCH codes as much as possible. RS codes are widely used and form an important class of algebraic codes. Towards the end, we will understand convolutional codes and the Viterbi decoding algorithm. Finally, we will get an idea of modern coding theory and iterative decoding methods that allow us to operate very close to capacity. A very good reference for channel coding is the book by Todd K. Moon [2].

References

[1] C. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, pp. 379–423, July 1948.

[2] T. K. Moon, *Error Correction Coding: Mathematical Methods and Algorithms*. John Wiley & Sons, Inc., 2005.