Analysis and simulation of a digital transmission system

Alessandro Trigolo

2023/2024

Index

Todo list

Part I

Analysis

1   Source data

The source data analysis provides a general overview of how the data are generated and how this will impact the encoding scheme. Specifically, the source data analysis is achieved by calculating two important values: the source entropy and the redundancy coefficient.

Source entropy

The source entropy H and the maximum source entropy H_{max}. The entropy of a sequence of symbols is a number that summarizes the randomness of the selection of the symbols in the source sequence. The more uncertain the symbols are, the higher the entropy is and the higher the information the symbols carry. The ideal entropy is when the source symbols are `1 0 1 0 1 0 ...` while the worst entropy is when all the symbols are `1` or `0`. Given a sequence S of N symbols, where each of them has its probability P_i to occur, the entropy of the sequence is:

$$H(S) = -\sum_{i=1}^{N} P_i \log_2 P_i$$

The entropy calculation can be simply achieved with the following MATLAB code. The only thing to note is that P is the probability vector that assigns to every symbol of the alphabet its probability.

```
% sum(P .* log2(P))
H = - dot(P, log2(P));
```

Secondly, in order to calculate the maximum entropy H_{max}, two conditions have to be met: all of the symbols have the same probability $P_i = \frac{1}{N}$ and, of course, they do not correlate one another. Consequently:

$$H_{max}(S) = -\sum_{i=1}^{N} \frac{1}{N} \log_2 \frac{1}{N} = \frac{1}{N} \sum_{i=1}^{N} \log_2 N = \log_2 N$$

Also in this case the MATLAB script to calculate the maximum entropy is trivial.

```
% Number of symbols in the alphabet
N  = length(P);

% Maximum source entropy
H_max = log2(N);
```

Redundancy coefficient

The redundancy coefficient ρ summarizes in a number how much additional information is present inside the sequence. Essentially, the lower the redundancy coefficient is, the better, because it means that the source entropy is very high. Mathematically, the coefficient ρ can be obtained as follows:

$$\rho = 1 - \frac{H}{H_{max}}$$

which translates into the following code snippet:

```matlab
% Calculate the redundancy coefficient 'rho'
source_redoundancy = 1 - H/H_max;
```