Data Transmission with Reliability Indication

Peter Hoeher
Information and Coding Theory Lab
University of Kiel
Kaiserstr. 2, D-24143 Kiel, Germany
Email: ph@techfak.uni-kiel.de

Abstract

In most digital communication systems, the information bits to be transmitted are protected against transmission errors. Typically, in the transmitter redundancy is added in a clever way by means of an encoder. In the receiver, a corresponding decoder is able to detect or to correct transmission errors by exploiting the redundancy. This technique is called channel coding. Applications of channel coding include digital cellular radio, data modems, wireless local loops, satellite systems, digital broadcasting systems, and even the compact disk.

Conventionally, the decoder does not output any reliability information about the decoded bits. Hence, it is not clear which of the decoded bits are reliable, if any. That knowledge would be important for subsequent processing stages, however.

Therefore, decoding with reliability information has become a popular area of research particularly in the last decade: the transmission power can be significantly reduced given the same transmission medium and a similar bit error rate performance.

In this tutorial, we firstly introduce the basic concept of encoding and decoding. Then, we illustrate a technique to obtain reliability information, before we finally feature some examples. We point out similarities with other research areas, for example classical analog transmission, equalization, digital filters, neuronal networks, etc.

1 Introduction

FIFTY years ago, Claude E. Shannon established the mathematical foundations for information transmission in a pioneering work [1]. He formulated the basic problem of reliable transmission of information in statistical terms, established basic limits on information processing, and gave birth to a new field that is now called information theory, covering

- source coding (data compression),
- channel coding (error detection and correction), and
- cryptography (secrecy coding).
Our focus is now on channel coding. To be specific, Shannon demonstrated that if the information rate from a source is less than the so-called channel capacity, then it is theoretically possible to achieve reliable transmission through the channel by appropriate channel coding. The channel capacity, which depends on power, bandwidth, and noise constraints, is the frontier in the field. For example, on a typical 3 kHz analog telephone line, the theoretical maximum information rate is about 30 kbit/s.

Unfortunately, information theory does not provide practical solutions to approach the limit. The ultimate goal is to reach the capacity limit as close as possible given a reasonable complexity in the transmitter and the receiver, and a maximum tolerable delay. Concerning the telephone line, today's technology supports data rates up to 28.8 kbit/s; this number is surprisingly close to the theoretical limit. (56 kbit/s modems rely on other constraints.)

State-of-the-art in channel coding are serial, parallel, and hybrid concatenated codes. Concatenated codes consist of more than one component code. They provide good distance properties (i.e., two codewords differ in many bits) yet using simple component codes, and motivate low-effort decoders. Besides designing good codes, the major challenge are low-complexity and low-delay decoder designs. These are theoretical and commercial challenges, respectively. Since mobility and information technology mutually become more important, the long-term significance of reliable data transmission is obvious.

2 Transmission System

Fig. 1 shows a block diagram of the coded transmission system under investigation.

\[ \begin{align*}
    u_i & \xrightarrow{\text{Encoder}} c_j \xrightarrow{\text{Channel}} + \xrightarrow{\text{Decoder}} \hat{u}_i = \hat{u}_i \cdot L_i \\
    0 \leq i \leq k - 1 & \quad 0 \leq j \leq n - 1
\end{align*} \]

Figure 1: Block diagram of a coded transmission system.

2.1 Encoder

We are given an uncoded data sequence \((u_0, u_1, \ldots, u_{k-1})\) with info bits \(u_i \in \{ \pm 1 \}\), where \(k\) is the length of the info sequence\(^1\). The encoder adds redundancy. The coded data sequence is denoted as \((c_0, c_1, \ldots, c_{n-1})\), where \(c_j \in \{ \pm 1 \}\) and \(n\) is the length of the coded sequence. The ratio \(R = k/n\) is called code rate. By definition, \(R < 1\). For example if \(R = 1/2\), which is a typical value in wireless systems, we have two coded bits per info bit.

\(^1\)Conventionally, the info bits are denoted as \(u_1 = 0\) and \(u_i = 1\), respectively. The advantage of defining \(u_i = +1\) and \(u_i = -1\) will become clear in the context of decoding with reliability information. Within this tutorial paper, we restrict ourself to binary i.i.d. data and binary codes.
2.2 Channel

The channel represents the transmission medium which is, by definition, non-ideal. Throughout, we assume that the disturbance is additive white Gaussian noise. This is a typical model for thermal noise at the receiver input. The noise samples, \( w_j \in \mathbb{R} \), are zero mean: 
\[
\mu = E[w_j] = 0.
\]

The variance is 
\[
\sigma^2 = E[w_j^2] = (2E_s/N_0)^{-1},
\]
where \( E_s \) is the energy per channel symbol (here: coded bit) and \( N_0 \) is the noise density. \( E_s/N_0 \) is called the signal/noise ratio (SNR) per channel symbol.

Since we transmit \( 1/R \) channel symbols per info bit, the energy per info bit is \( E_b = E_s/R \). Correspondingly, \( E_b/N_0 \) is the signal/noise ratio per info bit. This parameter is usually used to provide a fair comparison of different coded systems. According to Shannon, the minimum \( E_b/N_0 = \log_2 2 \approx -1.6 \text{ dB} \) if \( R \to 0 \). This theoretical limit is called Shannon bound.

2.3 Decoder

The task of the decoder is to recover the info sequence given the noisy channel symbols \( y_j \in \mathbb{R}, 0 \leq j \leq n - 1 \). By exploiting the redundancy of the coded sequence, transmission errors can be detected or even corrected. In the latter case, which is assumed in the following, the bit error rate after decoding is (typically much) less than the bit error rate on the channel.

The optimum decoder, in the sense of minimizing the bit error rate at the decoder output, computes

\[
\hat{u}_i = \hat{u}_i \cdot L_i = \log \frac{P(u_i = +1|(y_0, y_1, \ldots, y_{n-1}))}{P(u_i = -1|(y_0, y_1, \ldots, y_{n-1}))}, \quad 0 \leq i \leq k - 1.
\]

(1)

This computation can be done recursively [2].

A few points should be noted:

- The sign, \( \hat{u}_i \in \{\pm1\} \), is so-called hard output. Hard outputs are computed by any decoder.

- The magnitude, \( L_i \in \mathbb{R}_+ \), is called soft output. Soft outputs represent the reliability information.
  - For example, if \( P(u_i = +1|(y_0, y_1, \ldots, y_{n-1})) = P(u_i = -1|(y_0, y_1, \ldots, y_{n-1})) = 0.5 \), we obtain \( L_i = 0 \) and hence \( \hat{u}_i = 0 \). The hypotheses \( u_i = +1 \) and \( u_i = -1 \) are equally likely, so we obtain an “erasure”. This motivates our notation \( u_i \in \{\pm1\} \).
  - For example, if \( P(u_i = +1|(y_0, y_1, \ldots, y_{n-1})) = 10^{-6} \), we obtain \( \hat{u}_i = -1 \) and \( L_i \approx 6 \). Now the likelihood value, \( L_i \), indicates a reliable decision: On average, we have just 1 error in 1,000,000 info bits.

- The entire observations \( y_j, 0 \leq j \leq n - 1 \), are taken into account (as opposed to symbol-by-symbol decisions). Since we accept real-valued observations, \( y_j \), and since we deliver real-valued estimates, \( \hat{u}_i \), we have a soft-input soft-output decoder.

- Eqn. (1) represents a so-called maximum a posteriori decoder [2].

- Sub-optimal (but low effort) simplifications of (1) are proposed and studied in [3, 4], for example.
3 Experiments

Let us firstly consider a single code, and afterwards concatenated codes.

3.1 Single Code

A prominent class of channel codes, called *convolutional codes*, can be represented by a shift register. An example of a recursive systematic convolutional code is illustrated in Fig. 2.

![Recursive systematic convolutional code with 4 states (\(u_i \in \{\pm 1\}\)).](image)

Figure 2: Recursive systematic convolutional code with 4 states (\(u_i \in \{\pm 1\}\)).

Obviously, the code rate is \(R = 1/2\), since we output two coded bits, \(c_j\), per uncoded info bit, \(u_i\). The error correcting capabilities are mainly determined by the number of delay elements. In this example, we have 2 delay elements, corresponding to 4 states.

![Bit error rate using a \(R = 1/2\) recursive systematic convolutional code with 4 states and maximum a posteriori decoding. The uncoded performance (top curve) and the capacity bound (bottom curve) are featured as well.](image)

Figure 3: Bit error rate using a \(R = 1/2\) recursive systematic convolutional code with 4 states and maximum a posteriori decoding. The uncoded performance (top curve) and the capacity bound (bottom curve) are featured as well.

Given a white Gaussian noise channel and a maximum a posteriori decoder, the bit error rate of the coded scheme (obtained by Monte Carlo simulations) is plotted in Fig. 3. As a reference, the uncoded performance is shown in the same figure, as well as the capacity...
bound for $R = 1/2$ and binary signaling. (The two reference curves are analytical results.) Note from the figure that, at a bit error rate of $10^{-5}$, the coding gain is about 3.2 dB: The transmission power could be reduced by about 50% given the same bit error performance. However, we are still about 6.2 dB off the capacity bound. At high SNR, the coding gain is 4 dB.

Now, we are going to exploit the reliability information of the soft-input soft-output decoder. Recall that the variance of the channel, $\sigma^2$, corresponds to a signal/noise ratio of $E_s/N_0 = (2\sigma^2)^{-1}$. This parameter is dubbed $\text{SNR}_{\text{in}}$ in the following. If we assume that $u_i$ given $u$ is also Gaussian distributed, we may define $\text{SNR}_{\text{out}} = (2\sigma_{\text{out}}^2)^{-1}$, where $\sigma_{\text{out}}^2$ is the corresponding variance. In Fig. 4 we have plotted $\text{SNR}_{\text{out}}$ versus $\text{SNR}_{\text{in}}$ for the given scenario. Note that, for high SNR, the SNR improvement corresponds to the coding gain!

![SNR vs SNR](image.png)

Figure 4: $\text{SNR}_{\text{out}}$ vs. $\text{SNR}_{\text{in}}$ for $R = 1/2$ recursive systematic convolutional code with 4 states.

Fig. 4, which has firstly been investigated in [3], motivated numerous interesting analogies, interpretations, and further studies:

- Coding schemes with reliability information are related to classical analog modulation schemes such as frequency modulation (analog FM). This interpretation may be generalized to other uncoded or coded digital modulation formats (such as trellis-coded modulation), equalization, etc.

- The SNR improvement is related to nonlinear digital decimation filters, i.e., multi-rate systems. (The decimation factor is $1/R$.)

- This point of view motivated iterative decoding of concatenated codes [5, 6], see Section 3.2. These so-called turbo codes became perhaps the most important invention in coding theory over the last ten years.

- Soft-input soft-output techniques are related to fuzzy logic.

- Very recently, truly analog decoders were proposed [8, 9]. These time-continuous and value-continuous decoders may be seen as a special class of neuronal networks.
3.2 Turbo Codes

As mentioned in Section 1, concatenated codes are among the most powerful codes we are aware of in these days. The distance properties are good, and low cost decoders can be matched to the component codes; decoding the overall code would be extremely difficult.

The novel aspect of decoding concatenated codes is to pass reliability information between the component decoders. Thus, iterative decoding (which has been proposed by Gallager in the 60s) becomes extremely powerful. Corresponding schemes using convolutional codes and block codes were simultaneously proposed by Berrou et al. [5] and Lodge et al. [6]. Serial, parallel, and hybrid concatenations (“code networks”) [7], together with iterative processing, became known as turbo codes.

The efficiency of turbo codes is illustrated best with an example. We applied a parallel concatenation of two 4-state recursive systematic convolutional codes, see [5] for details. The length of the info sequence was chosen to be \( k = 10^5 \). The corresponding bit error results using maximum a posteriori component decoders are plotted in Fig. 5. Note that with 20 iterations the capacity bound can be reached as close as 0.6 dB given simple component codes and component decoders, i.e., given manageable complexity! Compared to Fig. 3, where the same component code is used, this is a remarkable improvement. The drawback is the decoding delay: Information theory tells us that channel capacity increases with the code length.

![Figure 5: Bit error rate for \( R = 1/2 \) turbo code with two 4-state component codes.](image)

4 Conclusions and Future Work

This tutorial highlights issues of decoding with reliability information. Interestingly, “soft information” (i) gives rise to numerous analogies which put channel decoding in a broader context, but (ii) also motivated practical encoding (“code networks”) and decoding techniques (“turbo processing”), which are currently state-of-the-art.
Open research topics include the analysis of soft-input soft-output decoders (including possible stability problems with iterative processing), low-effort and low-delay encoder and decoder designs, and analog implementations [8, 9].

References


