Library of decoders
Project proposal for SPCodingSchool

Marcelo Firer and João Strapasson

November 24, 2014

Let $\mathcal{P} = (\mathcal{X}, \mathcal{Y}, P)$ be a channel with output set $\mathcal{Y}$ and input set $\mathcal{X} \subseteq \mathcal{Y}$ and transition matrix P_{xy}, where is defined by the conditional probabilities.

$$P_{xy} = \Pr(x \text{ sent} \mid y \text{ received})$$

In many cases $\mathcal{X}$ and $\mathcal{Y}$ may be defined as an n-fold product of a given alphabet ($\mathcal{X} = \mathbb{F}_q^n$ a vectors space over a finite field, for example), but this is irrelevant to this project.

A code is just a non-empty subset $C \subseteq \mathcal{X}$ and an element of C is called a codeword or message. In general, when a message $x \in C$ is transmitted through the channel, a message $y \in \mathcal{Y}$ is received and a *decoder* for C is a map $a : \mathcal{Y} \mapsto C$ that tries to recover x, so that we may assume that $a(x) = x$ whenever $x \in C$. We may consider a decoder to list a set of possible codewords, that is, we may consider a decoder to be a map $a : \mathcal{Y} \to \mathbb{P}^*(C)$ where $\mathbb{P}^*(C) = \{A | \emptyset \neq A \subseteq C\}$ is the set of non empty subsets of C. We say that a decoder determines a unique decision for y if $a(y)$ is a set consisting of a unique element and a is a unique-decision decoder if it is so for every $y \in \mathcal{Y}$. An *universal decoder* for the channel $\mathcal{P} = (\mathcal{X}, \mathcal{Y}, P)$ is a map $\alpha : \mathbb{P}^*(\mathcal{X}) \to F(\mathcal{Y}, \mathbb{P}^*(\mathcal{X}))$ that associates to each code $C \in \mathbb{P}^*(\mathcal{X})$ a decoder of $\alpha(C) = a_C : \mathcal{Y} \to \mathbb{P}^*(C) \subset \mathbb{P}^*(\mathcal{X})$.

Of course, an universal decoder can be described by making an (huge) list of decoders $\alpha(C_1), ..., \alpha(C_{2^{|\mathcal{X}|}})$ and for each of those $2^{|\mathcal{X}|}$ decoders listing all the $2^{|\mathcal{Y}|}$ decisions $a_{C_i}(y_1), ..., a_{C_i}(y_{2^{|\mathcal{Y}|}})$. However, there are two sources of universal decoders that can be easily described (not implemented), the *Maximal Likelihood* (ML) decoders and the *Nearest Neighbor* (NN) decoders.

The ML decoder is a set of probabilistic decision criteria determined by the channel|: given $C \in \mathbb{P}^*(\mathcal{X})$, the ML decoder $a_C^{ML} : \mathcal{Y} \to \mathbb{P}^*(C)$ is defined by

$$a_C^{ML}(y) = \arg\max_{x \in C} \Pr(x \text{ sent} \mid y \text{ received})$$

where

$$\begin{aligned} &\arg\max_{x \in C} \Pr(x \text{ sent} \mid y \text{ rec}) \\ &:= \{c \in C | \Pr(c \text{ sent} \mid y \text{ rec}) \geq \Pr(x \text{ sent} \mid y \text{ rec}), \forall x \in C\} \end{aligned}$$

is the set of most probable codewords to be sent once that y was the received message.

When $\mathcal{Y}$ is endowed with a metric $d : \mathcal{Y} \times \mathcal{Y} \to \mathbb{R}^+$, we can defined an NN decoder, as a map that associates to a given $y \in \mathcal{Y}$ the set of codewords that minimizes the distance to y, that is, $a_C^{NN} : \mathcal{Y} \to \mathbb{P}^*(C)$ is defined by

$$a_C^{NN}(y) = \arg\min_{x \in C} d(x, y)$$

where

$$\arg\min_{x \in C} d(x, y) := \{c \in C | d(c, y) \leq d(x, y), \forall x \in C\}.$$

Despite the fact that ML decoders and NN decoders are easy to describe, they are in general extremely difficult to be implemented. It the most common setting where $\mathcal{X} = \mathbb{F}_q^n$, given a code $C \subset \mathcal{X} \subseteq \mathcal{Y}$ and a decoder (either ML or NN) $a : \mathcal{Y} \to \mathcal{C}$, a search algorithm involves a list of the size of $\mathcal{Y}$, that increases exponentially with n.

Small values of n are suitable for situations with strong constrains at block length (see, for example, [2] or the introduction in [1]) and for research proposes, since actually listing decoders is essential for computing many of the important invariants (such as error probability).

MAIN GOAL: The main goal of this project is to create a library of ML and NN decoders for use with software for numerical analysis, such as Mathematica, Maple, Matlab, etc.

How to approach the problem:

If we consider $\mathcal{X} = \mathcal{Y} = \mathbb{F}_q^n$, linear codes is a family that may be used to attain the channel capacity and simplifies calculations. We say that a channel $\mathcal{P} = (\mathcal{X}, \mathcal{Y}, P)$ is *invariant by translations* if

$$\Pr((x+z) \text{ sent} \mid (y+z) \text{ received}) = \Pr(x \text{ sent} \mid y \text{ received})$$

for all $x, y, z \in \mathbb{F}_q^n$. Similarly, a metric $d : \mathbb{F}_q^n \times \mathbb{F}_q^n \to \mathbb{R}$ is said to be *invariant by translations* if

$$d(x+z, y+z) = d(x, y)$$

for all $x, y, z \in \mathbb{F}_q^n$. There are plenty of metrics that are invariant by translations, and those includes the Hamming metric, Lee metric, Poset metrics and any metric determined by a weight (see for example Section 16 in [3]). For those channels and metrics decoding of linear codes may be done using syndromes. Indeed, let $C \subset \mathbb{F}_q^n$ be a k-dimensional linear code and let $C, x_1 + C, ..., x_{N-1} + C$ be the N distinct lateral classes (cosets), where $N = 2^{n-k}$. We assume that each x_i is a representative of minimal weight of the class (*syndrome leader*), that is, by denoting $w(x) = d(x, 0)$ we have that

$$w(x_i) \leq w(x_i + c), \ \forall c \in C.$$

Considering a metric that is invariant by translation, given $y \in x_i + C$ we have that $y - x_i \in C$ and

$$d(y, y - x_i) \leq d(y, c), \ \forall c \in C$$

so, in order to decode y we need to determine in which of the N distinct coset y is contained. Since the code is linear, by considering a parity check matrix H, since $H(x_i + c)^T = H(x_i)^T$ we actually need to search in a table containing $N = 2^{n-k}$ elements $0, H(x_1)^T, ..., H(x_{N-1})^T$ (instead of the 2^n elements of $\mathbb{F}_q^n$).

We remark that determining the syndrome leaders is by itself a difficult task (in general). However,if we know that the packing radius R of the code C, we do know that all elements $x \in \mathbb{F}_q^n$ such that $w(x) \leq R$ are syndrome leader. We remark that for general metrics the packing radius is not necessarily determined by the minimal distance and in its full generality, it is by itself a very hard task, as exposed in [4].

Similar considerations arises when considering a channel invariant by translation and defining a syndrome leader to be a vector x_i such that

$$\Pr(x_i \text{ received} \mid 0 \text{ sent}) \geq \Pr((x_i + c) \text{ received} \mid 0 \text{ sent}), \ \forall c \in C.$$

Prerequisites and suggested readings

Reasonable knowledge of Mathematica or Matlab.

References

[1] S. M. Moser - *Weak Flip Codes and Applications to Optimal Code Design on the Binary Erasure Channel* - research report available at http://moser-isi.ethz.ch/docs/papers/smos-2013-6.pdf, 2013.

[2] Po-Ning Chen; Hsuan-Yin Lin; Moser, S.M. - *Optimal Ultrasmall Block-Codes for Binary Discrete Memoryless Channels* - Information Theory, IEEE Transactions on , vol.59, no.11, pp.7346,7378, 2013.

[3] Michel Marie Deza and Elena Deza, *Encyclopedia of distances*, second ed., Springer, Heidelberg, 2013.

[4] Rafael Gregorio Lucas D'Oliveira, Marcelo Firer, *The Packing Radius of a Code and Partitioning Problems: the Case for Poset Metrics,* arXiv:1301.5915 [cs.IT], 2013.