Yugoslav Journal of Operations Research
8 (1998), Number 1, 169-175
Jovan Petrić memorial volume

SOFT DECISION DECODING OF BINARY LINEAR BLOCK CODES USING THE BFGS-METHOD

Klaus RITTER
Technische Universität München, Zentrum Mathematik,
Arcisstr. 21, 80290 München, Deutschland

Stefan SCHÄFFLER, Thomas F. STURM
SIEMENS AG, Zentralabteilung Technik ZT PP 2S,
81730 München, Deutschland

Dedicated to the memory of Professor Jovan Petrić

Abstract: This paper presents a new method for soft decision decoding of binary linear block codes. An efficient algorithm for the numerical computation of soft outputs given by an AWGN channel is developed and investigated based on a special nonlinear regression model. The minimization of the sum of squared errors in the considered regression model leads to an unconstrained nonlinear optimization problem with continuously differentiable objective function, which is numerically solved by the BFGS-method. Numerical results using BCH codes are summarized.

Keywords: Digital communications, soft decision decoding, binary block codes, nonlinear regression, unconstrained optimization, BFGS-method.

1. INTRODUCTION

The theory of information founded by Claude Shannon in 1948 describes the possibilities and limits of coding theory. Since this description is not constructive in general, one has to develop efficient codes and decoding algorithms which make use of the whole information of the underlying communication system.

In order to correct errors caused by the transmission of k information bits $\mathbf{u} \in \{\pm 1\}^k$ through an AWGN (Additive Gaussian White Noise) channel each code word

u is enlarged by $n-k$ security bits (redundancy) $c_i \in \{\pm 1\}$, $i = k+1,\ldots,n$. An algebraic field structure of the set $\{\pm 1\}$ given by the commutative operators $\oplus$ and $\otimes$ with

$$\begin{aligned} -1 \oplus -1 &= 1 \\ 1 \oplus 1 &= 1 \\ 1 \oplus -1 &= -1 \\ -1 \otimes -1 &= -1 \\ 1 \otimes -1 &= 1 \\ 1 \otimes 1 &= 1 \end{aligned}$$

is used for the computation of the security bits. Each security bit c_i is defined by a fixed set $J_i \subset \{1,\ldots,k\}$:

$$c_i = \bigoplus_{j \in J_i} u_j, \quad i = k+1,\ldots,n \tag{1.1}$$

The set

$$C = \{\mathbf{c} \in \{\pm 1\}^n : c_i = u_i, i = 1,\ldots,k, \mathbf{u} \in \{\pm 1\}^k, c_i = \bigoplus_{j \in J_i} u_j, i = k+1,\ldots,n\}$$

of all possible code words depending on k, n, and $J_{k+1},\ldots,J_n$ is called a (n, k) (systematic) binary linear block code. The optimal design of binary linear block codes is a main topic in information theory and in coding theory (see, e.g. [1], [5], and [7]).

Using a digital communication system with AWGN channel, the properties of this communication system are mainly determined by the signal to noise ratio $\frac{E_b}{N_0}$ which represents the proportion between signal transmission energy for one information bit and channel noise energy. The information at the receiver after the transmission of a code word **c** trough an AWGN channel with signal to noise ratio $\frac{E_b}{N_0}$ can be interpreted as a realization $\mathbf{y} \in \mathbf{R}^n$ of a $\mathcal{N}(\mathbf{c}, \frac{N_0 n}{2E_b k}\mathbf{I}_n)$ Gaussian distributed random vector with mean value **c** and covariance matrix $\frac{N_0 n}{2E_b k}\mathbf{I}_n$, where **c** is the transmitted code word of a (n, k) binary linear block code and where $\mathbf{I}_n$ denotes the n-dimensional identity matrix. A decoding algorithm has to estimate the first k components of **c** based on the known code, the known signal to noise ratio, and based

on the received information $\mathbf{y}$. This type of decoding is called soft decision decoding in contrast to hard decision decoding where sign (y_i) is used instead of y_i for all $i = 1,...,n$.

For a large class of applications (see [4]) it is necessary to compute for each estimated information bit a security measure which quantifies the reliability of the decision made by the decoding algorithm. Therefore, in this paper we investigate a new decoding algorithm which computes a vector $\tilde{\mathbf{u}} \in \mathbf{R}^k$ based on an AWGN channel with known code, known signal to noise ratio, and based on the received information $\mathbf{y}$. An estimation for the ith information bit is given by $\operatorname{sign}(\tilde{u}_i)$. The absolute value of $\tilde{u}_i$ indicates the reliability of this decision. If $\tilde{u}_i = 0$ then both possible values of the ith information bit have the same probability. Hance, the decision is purely random without reliability in this case. The vector $\tilde{\mathbf{u}}$ can be interpreted as a numerical approximation of a special security measure which is called soft output. The approximation of soft outputs is based on nonlinear regression models which are investigated in the next section. The minimization of the sum of squared errors in these regression models leads to unconstrained nonlinear optimization problems with continuously differentiable objective functions, which are numerically solved by the BFGS-method. Finally, numerical results using BCH codes are presented.

2. AN APPROXIMATION OF SOFT OUTPUTS

In thi section, we consider a digital communication system with AWGN channel, known signal to noise ratio $\frac{E_b}{N_0}$, known received information $\mathbf{y}$, and fixed binary linear block code represented by

$$C = \{\mathbf{c} \in \{\pm 1\}^n;\ c_i = u_i,\ i = 1,\ldots,k,\ \mathbf{u} \in \{\pm 1\}^k,\ c_i = \bigoplus_{j \in J_i} u_j,\ i = k+1,\ldots,n\}$$

A stochastic analysis given in [4] and [8] shows the existence of an optimal security measure for decoding the first k components of a transmitted code word $\mathbf{c}$. This security measure is represented by a vector $\mathbf{w} \in \mathbf{R}^k$ with components defined by

$$w_i = \ln\left(\frac{\displaystyle\sum_{\substack{\mathbf{v}\in C\\ v_i=+1}} \exp\left(-\frac{(\mathbf{y}-\mathbf{v})^T(\mathbf{y}-\mathbf{v})}{\frac{N_0 n}{E_b k}}\right)}{\displaystyle\sum_{\substack{\mathbf{v}\in C\\ v_i=-1}} \exp\left(-\frac{(\mathbf{y}-\mathbf{v})^T(\mathbf{y}-\mathbf{v})}{\frac{N_0 n}{E_b k}}\right)}\right), \quad i = 1,\ldots,k \tag{2.2}$$

An estimation of the ith information bit is given by $\operatorname{sign}(w_i)$. The absolute value of w_i indicates the reliability of this decision. The vector $\mathbf{w}$ is called soft output. Unfortunately, the number of operations for the computation of one component of $\mathbf{w}$ is given by $\min(2^k, 2^{n-k})$. Therefore, it is necessary to find numerical approximations for the soft outputs for a large class of codes.

A profound analysis of the underlying communication system shows that the components of $\mathbf{w}$ can be expressed by a nonlinear regression model

$$\frac{4E_b k}{N_0 n}\mathbf{y} = \begin{pmatrix} w_1 \\ \vdots \\ w_k \\ \ln\left(\dfrac{1+\displaystyle\prod_{j\in J_{k+1}}\dfrac{\exp(w_j)-1}{\exp(w_j)+1}}{1-\displaystyle\prod_{j\in J_{k+1}}\dfrac{\exp(w_j)-1}{\exp(w_j)+1}}\right) \\ \vdots \\ \ln\left(\dfrac{1+\displaystyle\prod_{j\in J_n}\dfrac{\exp(w_j)-1}{\exp(w_j)+1}}{1-\displaystyle\prod_{j\in J_n}\dfrac{\exp(w_j)-1}{\exp(w_j)+1}}\right) \end{pmatrix} + \mathbf{e}, \tag{2.3}$$

where the vector $\mathbf{e}$ of errors is zero if the noise energy of the AWGN channel is zero (i.e., the covariance matrix of the Gaussian distribution is the zero matrix).

The minimization of the sum of the squared components of $\mathbf{e}$ leads to the following unconstrained optimization problem

$$\min_{\mathbf{w}\in\mathbf{R}^k}\left\{\mathbf{e}(\mathbf{w})^T\mathbf{e}(\mathbf{w})\right\} =$$

$$\min_{\mathbf{w}\in\mathbf{R}^k}\Bigg\{\underbrace{\sum_{i=1}^{k}\left(w_i-\frac{4E_bk}{N_0n}y_i\right)^2+\sum_{i=k+1}^{n}\left(\ln\left(\frac{1+\displaystyle\prod_{j\in J_i}\frac{\exp(w_j)-1}{\exp(w_j)+1}}{1-\displaystyle\prod_{j\in J_i}\frac{\exp(w_j)-1}{\exp(w_j)+1}}\right)-\frac{4E_bk}{N_0n}y_i\right)^2}_{\text{objective function } f}\Bigg\}.$$

The existence of a global minimizer of the objective function f is guaranteed by the fact that f is bounded below by a paraboloid:

$$f(\mathbf{w}) \geq \sum_{i=1}^{k} \left(w_i - \frac{4E_b k}{N_0 n} y_i \right)^2 \tag{2.4}$$

Since the computation of the Hessian matrix at any chosen point is very expensive, the BFGS-method is used for the numerical minimization of f in the next section.

3. NUMERICAL RESULTS

The approximation of soft outputs using nonlinear regression models leads to unconstrained optimization problems which we solved numerically using the BFGS-method with analytically computed first derivatives. This method is a reliable optimization procedure which is widely used in nonlinear programming (see [2] and [6]).

BFGS-method

Step 0: (Initialization)

Choose $\mathbf{w}_0$, a positive definite matrix $\mathbf{H}_0$, and γ_1, γ_2 with:

$$0 < \gamma_1 < 0.5 \,,\ \gamma_1 < \gamma_2 < 1 \,.$$

Compute $\nabla f(\mathbf{w}_0)$.

If $\| \nabla f(\mathbf{w}_0) \|_2 \leq 10^{-4}$ then STOP; else: $j := 0$, go to step 1.

Step 1: (Computation of a search direction)

Compute

$$\mathbf{s}_j := \mathbf{H}_j \nabla f(\mathbf{w}_j) \,,$$

go to step 2.

Step 2: (Computation of a step size)

Compute σ_j with:

$$\nabla f(\mathbf{w}_j - \sigma_j \mathbf{s}_j)^T \mathbf{s}_j \leq \gamma_2 \nabla f(\mathbf{w}_j)^T \mathbf{s}_j \,,$$

$$f(\mathbf{w}_j - \sigma_j \mathbf{s}_j) \leq f(\mathbf{w}_j) - \gamma_1 \sigma_j \nabla f(\mathbf{w}_j)^T \mathbf{s}_j \,,$$

$\sigma_j = 1$, if possible.

Compute

$$\mathbf{w}_{j+1} = \mathbf{w}_j - \sigma_j \mathbf{s}_j .$$

If $\| \nabla f(\mathbf{w}_{j+1}) \|_2 \leq 10^{-4}$ then STOP; else: go to step 3.

Step 3: (Computation of $\mathbf{H}_{j+1}$)

With

$$\mathbf{d}_j := \frac{\nabla f(\mathbf{w}_j) - \nabla f(\mathbf{w}_{j+1})}{\| \sigma_j \mathbf{s}_j \|_2}, \quad \mathbf{p}_j := \frac{\mathbf{s}_j}{\| \mathbf{s}_j \|_2}$$

compute

$$\mathbf{H}_{j+1} = \mathbf{H}_j + \frac{\mathbf{d}_j^T \mathbf{p}_j + \mathbf{d}_j^T \mathbf{H}_j \mathbf{d}_j}{(\mathbf{d}_j^T \mathbf{p}_j)^2} \mathbf{p}_j \mathbf{p}_j^T - \frac{\mathbf{p}_j \mathbf{d}_j^T \mathbf{H}_j + \mathbf{H}_j \mathbf{d}_j \mathbf{p}_j^T}{\mathbf{d}_j^T \mathbf{p}_j} .$$

$j := j + 1$,
go to step 1.

In Table 1 we provide numerical results for BCH(n, k) codes (Bose Chaudhuri Hocquenghem), which are powerful binary linear block codes, with several signal to noise ratios. The probability of bit errors is computed as follows: For each code as many code words are decoded as are necessary to obtain at least 150 bit errors. Altogether, 1300 optimization problems have been solved using the BFGS-method. The results are comparable to results obtained by hard decision decoding of the same codes with about 1 dB more signal to noise ratio (see [3]). Therefore, a 20 percent reduction of the transmitting energy is achieved applying our new method instead of hard decision decoding at equal transmission quality.

Table 1: Probability of bit errors for several codes and several signal to noise ratios with the new method

$\frac{E_b}{N_0}$ in dB	BCH(31, 21)	BCH(63, 30)	BCH(127, 99)
0	1.07e-1	1.84e-1	9.44e-2
0.5	9.45e-2	1.51e-1	9.80e-2
1	8.23e-2	1.35e-1	8.19e-2
1.5	6.67e-2	1.17e-1	6.52e-2
2	4.91e-2	1.10e-1	5.92e-2
2.5	3.70e-2	8.27e-2	5.00e-2
3	2.96e-2	7.65e-2	4.26e-2

REFERENCES

[1] Ash, R.B., *Information Theory*, Interscience, New York, 1965.

[2] Fletcher, R., *Practical Methods of Optimization*, Wiley, Chichester, 1993.

[3] Friedrichs, B., *Kanalcodierung*, Springer, Heidelberg, 1996.

[4] Hagenauer, J., Offer, E., and Papke, L., "Iterative decoding of binary block and convolutional codes", *IEEE Transactions on Information Theory*, 42 (1996).

[5] Heise, W., and Quattrocchi, P., *Informations- und Codierungstheorie*, Springer, Heidelberg, 1995.

[6] Luenberger, D.G., *Linear and Nonlinear Programming*, Addison-Wesley, Massachusetts, 1984.

[7] Proakis, J.G., *Digital Communications*, McGraw-Hill, New York, 1995.

[8] Schäffler, S., *Decodierung binärer linearer Blockcodes durch globale Optimierung*, Roderer, Regensburg, 1997.