VARIABLE SELECTION METHOD USING STATISTICAL SENSITIVITY ANALYSIS FOR RADIAL BASIS FUNCTION NETWORKS: APPLICATION TO ADAPTIVE CHANNEL EQUALISATION

F.Albu [1], B. Dorizzi [2], J. C. M. Mota [2,3]

[1] "POLITEHNICA" University of Bucharest, 1-3 Bd. Iuliu Maniu,.felix@pns.comm.pub.ro
[2] INT, EPH, 9 Rue Charles Fourier, 91011 Evry , France, Bernadette.Dorizzi@int-evry.fr
[3] Federal University of Ceara, Electrical Eng.Dept., CP:6001, 60455-760, Ceará, Brazil, mota@olympe.int-evry.fr

***Abstract*: This paper investigates the application of Radial Basis Functions Networks (RBFN) to the adaptive channel equalization of a bipolar signal passed through a dispersive channel in the presence of additive noise. The computational requirement to implement the optimal Bayesian symbol-decision equalizer using RBFN [1] can be very high as the optimal Bayesian solution requires a large number of centers. Statistical Sensitivity Analysis [3, 6] is proposed for selection of an appropriate subset of input variables which finally can lead to a more parsimonious RBF equalizer structure. Our simulation results show that this analysis method can provide a good compromise between RBFN complexity and equalization performance.**

1. Introduction

Channel equalization is a technique employed to combat the effects of intersymbol interference and noise which corrupt the transmission of signals across a communication channel (Fig. 1). The equalization objective is to reconstruct the transmitted sequence with the minimum error probability , i. e. : $\tilde{s}(t) = s(t-d)$, where d is the delay. The channel is usually modeled by a FIR filter with the following transfer function : $H(z) = \sum_{i=0}^{n_h-1} h_i z^{-i}$, where h_i are the channel impulse response components and n_h is his length. In our study the transmitted symbol $s(t)$ is taken from the data set {±1}; it forms an i.i.d. sequence, and $e(t)$ is an additive white Gaussian noise with zero mean and variance σ_e^2.

Classically, linear equalizer are considered for this task, but usually it reinforces the noise and it ignores the fact that $s(t)$ came from data set. A non-linear equalizer is prefered, particularly in radio mobil systems designs.

It is well-known [7] that a non-linear block detection equalization based on the principle of Maximum Likelihood Sequence Estimator will provide the best equalization performance when the channel is completely known. Its implementation complexity is one of the main reasons for using other non-linear symbol-decision class equalizers with simpler implementations but poorer performances. In this context, the communications community has recognized the Bayesian symbol-decision class equalizers as optimal solutions which deals with the equalization problem as a classification one[1].

Let us consider now RBFN [1]. It is a two layers network comprising a hidden layer and an output layer, and it has been shown to be capable of universal approximation [4]. The hidden layer contains n neurons which calculate the Euclidian distance between a center vector $\mathbf{c}_i$ and an input vector $\mathbf{y} = [y(t)\ y(t-1)\ \ldots\ y(t-m+1)]^t$, where in Fig.2: $y(t-j+1) = y_j$, $j \in [1;m]$. The result is passed through a nonlinear function Φ_i to generate the hidden node output. Functions Φ_i normally are chosen to be Gaussian: $\Phi_i = \exp(-\|\mathbf{y} - \mathbf{c}_i\|^2/\gamma_i^2)$, where γ_i is called the width. The output layer is computed by a weighted linear combination of the n neurons of the hidden layer. The overall response is a mapping: $f(\mathbf{y}) = \sum_{i=1}^{n} w_i \Phi_i$, where w_i are the weights.

It has been shown [2] that RBFN realize an implementation of the optimal Bayesian equalizer if the channel is known and the parameters of the network are well chosen (i. e., the number of hidden neurons n is equal to the number of desired channel states $\bar{y}(t)$: $n = 2^{m+n_h-1}$; the RBFN centers are placed at desired channel states vectors: $\mathbf{c}_i = [\bar{y}_i(t)\ \bar{y}_i(t-1) \ldots \bar{y}_i(t-m+1)]^t$; the weights are chosen in the data set: {±1}, and $\gamma_i^2 = 2\sigma_e^2$, $i \in [1;n]$). We remark that the RBFN structure can be complex when m and n_h are large.

In our work, the training of RBFN was done using a two-steps approach: in the first step a supervised k-means clustering procedure [2] was used to search and optimise the location of the centers and the widths were set at $r_i = r = \dfrac{d_m}{\sqrt{2 \cdot n}}$, where $i \in [1;n]$ and d_m is the maximum distance between the chosen centers. In the second step, the weights were trained using the least mean squares (LMS) algorithm. In this case, the RBFN equalization structure is called full RBF equalizer in this work. The signal error: $s_e(t) = \bar{s}(t) - s(t-d)$, is used in both the supervised clustering procedure and the LMS algorithm adaptive one.

2. Variable Selection Method using Statistical Sensitivity Analysis (VS-SSA)

The primary aim of this analysis is the selection of an appropriate subset of input variables in order to estimate the more performent and parsimonious RBFN structure. The method is based on the analysis of the partial derivatives of the RBF output $f(\mathbf{y}) = \bar{s}(t)$ with regards to its inputs y_j. Of course these quantities are random variables which have to be measured statistically, on the learning database. The analysis of the distributions of the sensitivities on the training set allows the selection of a candidate subset of irrelevant input variables. Different graphs enable us to make the selection of the relevant variables. We analyse the plot of the 95% quantile of the absolute value of the derivatives for each y_i. The quantiles are normalised to one because we are interested in the relative influence of each variable. If this influence is close to zero, the corresponding variable is considered as neglectible. The 95% are used to make the criterion more robust, although it will capture less variability of the derivatives. We also analize the mean versus the standard deviation (std) of each derivative and if the corresponding point is close to zero, the corresponding variable is considered as neglectible. Unfortunately, sometimes we have a fuzzy zone where it is hard to decide whether a variable is important or not. In those cases we will select the most important variables, trying to discard only the clearly irrelevant ones. We then re-estimate a RBFN structure using the selected input variables to compute more precise derivatives. Unfortunatelly the VS-SSA method has some drawbacks. Sometimes, the number of input variables will confuse the derivatives, because the model is over-parametrized. Even in those cases where the decision gets harder because all the variables tend to be important, we can determine the most and the least important inputs. In the presence of high noise and small amount of data, the problem gets tougher, because it is hard to estimate a good RBFN structure.

3. Simulation Results

The RBFN equalizer structures were analysed in three nonminimum phase channels. We used a number of samples for RBF training equal to the number of centers multiplied by 10 and a validation dataset to stop learning [3, 6]. The algorithms step-size parameters were optimized in all simulations. We observed that the convergence of the supervised k-means clustering procedure is usually attained. In all VS-SSA studies, the delay d was fixed to 1. The examples illustrate the good properties of the VS-SSA method for reducing the (m-n-1) RBFN complexity by eliminating some inputs and centers.

The first channel considered is: $H_1(z) = 1 + 0.729z^{-3}$. We used a 6-512-1 RBFN structure, with inputs $y(t), y(t-1), y(t-2), \ldots, y(t-5)$. The SNR (Signal-Noise Ratio) was 17 dB, and in this case, the explanatory variables $y(t-1)$ and $y(t-4)$ clearly pop up (Fig 3 a,b). Using only these inputs, we retrained the network as explained above and we were able to set to zero some weights which were very close to zero by using a threshold value equal to 1 percent of the maximum value of the weights. The number of pruned centers depends on the number of pruned inputs. In this case the number of centers can be reduced from 512 to 32 (Fig. 3c). The performance of the parsimonious RBFN's structure (reduced RBF equalizer) closely matches the full RBF equalizer performance in terms of $BER_L = \log_{10}BER$ (BER=Bit Error Rate), for 200000 test samples (Table 1). This result shows the robustness of the RBFN equalizer structure.

The second channel used in our simulations was a discrete microwave channel modeled as a FIR filter, where only three components are selected based on a maximum peak distortion criterium. His transfer function is: $H_2(z) = -0.0875 + 0.7901z^{-1} - 0.5989z^{-2}$. This channel model is obtained by sampling the analog two-rays propagation model [5]. We used a roll-off parameter equal to 0.3 in cosinus-raised filter system and a transmission rate equal to 24 Mbit/sec. Phase-offset is not considered and the sampling optimum epoch is used [5]. We designed a 9-2048-1 RBFN structure, with inputs $y(t)$, $y(t-1)$, $y(t-2)$, …, $y(t-8)$. In this case, as the distorsions are more important, the selection is more difficult. We discarded the last three inputs and the performance is acceptable for 17 dB (Fig. 4). The BER_L of the optimal Bayesian equalizer, the full RBF equalizer and the order-6 Wiener filter equalizer are plotted for comparison (Fig. 4c). We can see the superiority of nonlinear techniques and that the RBF equalizer achieves the optimal performance (Bayesian equalizer) independently of the SNR value.

We observed a damped repetition of statistical influence of the important inputs over the irrelevant ones especially when the transfer function of the channel has consecutive significant coefficients.

4. Conclusion and Perspectives

In this paper, we have proposed the Variable Selection Method using the Statistical Sensitivity Analysis for selection of an appropriate input variables subset which can lead to a parsimonious RBF equalization structure, and can reduce its complexity by diminution of the number of centers without significant degradation in equalization performance. Our future work will be to extend this research for more complex modulations, to investigate other clustering algorithms and radio mobile channel applications.

Acknowledgement

This work was partially supported by CNPq/Brazil. The authors would like to thank T. Czernichow for the

This work was partially supported by CNPq/Brazil. The authors would like to thank T. Czernichow for the
software toolbox and F. Jacquet for his useful comments.

References

[1] B.Mulgrew : “Applying Radial Basis Functions”, *IEEE Signal Processing Magazine*, March 1996.

[2] S. Chen, B. Mulgrew, P. M. Grant, “A clustering technique for digital communications channel equalization using Radial Basis Function networks”, *IEEE Transactions on Neural Networks*, Vol 4, July 1993.

[3] T. Czernichow, Application des réseaux neuronaux à la prediction de séries temporelles, *PhD Thesis*(in French), Université Paris VI, 1996.

[4] S. Haykin, Neural network : A Comprehensive Foundation, *Prentice Hall*, 1994.

[5] J.C.M.Mota, Equalização Adaptativa em Sistemas Rádios-Digitais: Das Técnicas em Freqüência à Desconvolução Autodidata, *PhD Thesis*(in Portuguese), UNICAMP/ Campinas, São Paulo, Brazil, August/1992.

[6] B. Dorizzi, G. Pellieux, F. Jacquet, T. Czernichow, A. Muñoz, “Selecting the relevant variables to forecast the French T-bond”, *Chemical Bank*, Mars 1996.

[7] J.G. Proakis Digital Communications, *McGraw- Hill*,1984

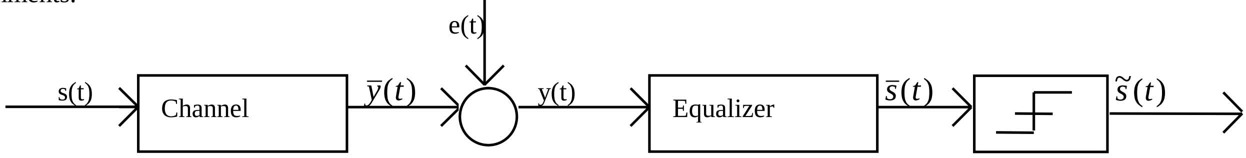

Fig. 1- Discrete-time model of a data transmission system.

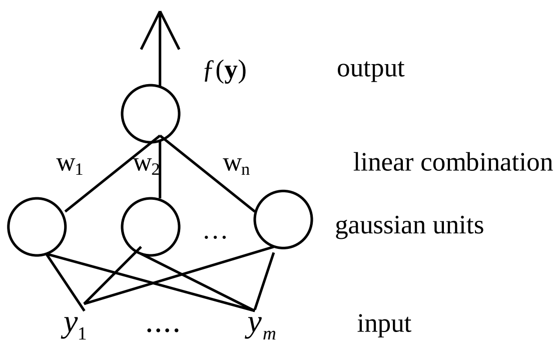

Fig. 2 - Radial Basis Function Network (RBFN).

	Optimal Bayesian Equalizer	Full RBF Equalizer(n=512)	Reduced RBF Equalizer(n=32)
BER_L	- 3.08	-3.06	-3.03

Table 1: $H_1(z)$ - BER_L of: optimal Bayesian equalizer (m=6), full RBF equalizer (m=6, n=512) and reduced RBF equalizer (m=2, n=32).

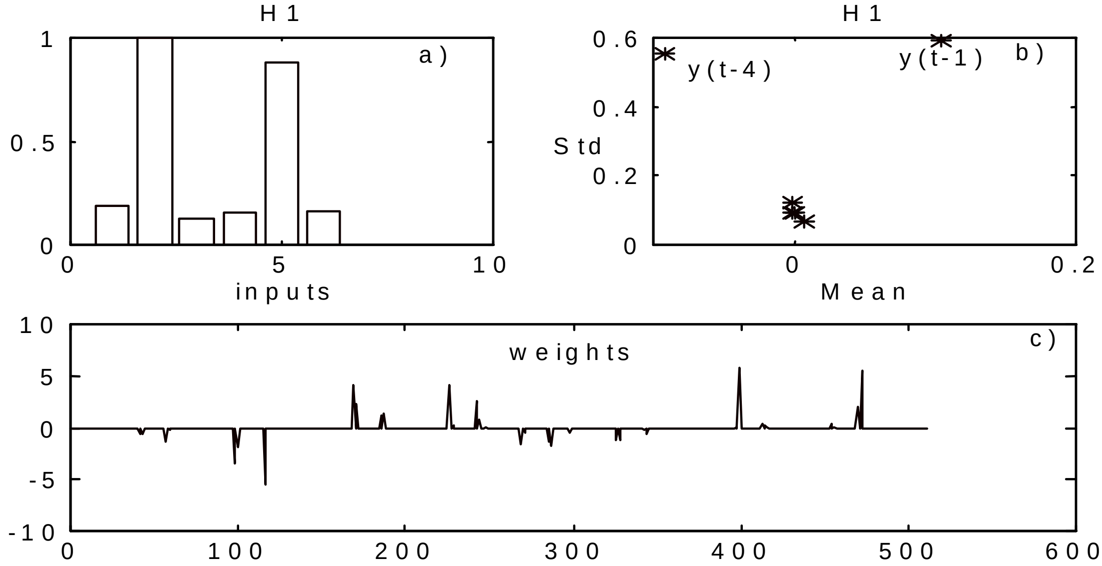

Fig. 3 - (a) 95% normalized quantile of the absolute derivatives. (b) Mean$\left(d\bar{s}/dy\right)$ x std$\left(d\bar{s}/dy\right)$.
(c) Weights of the pruned RBFN, SNR=17 dB.

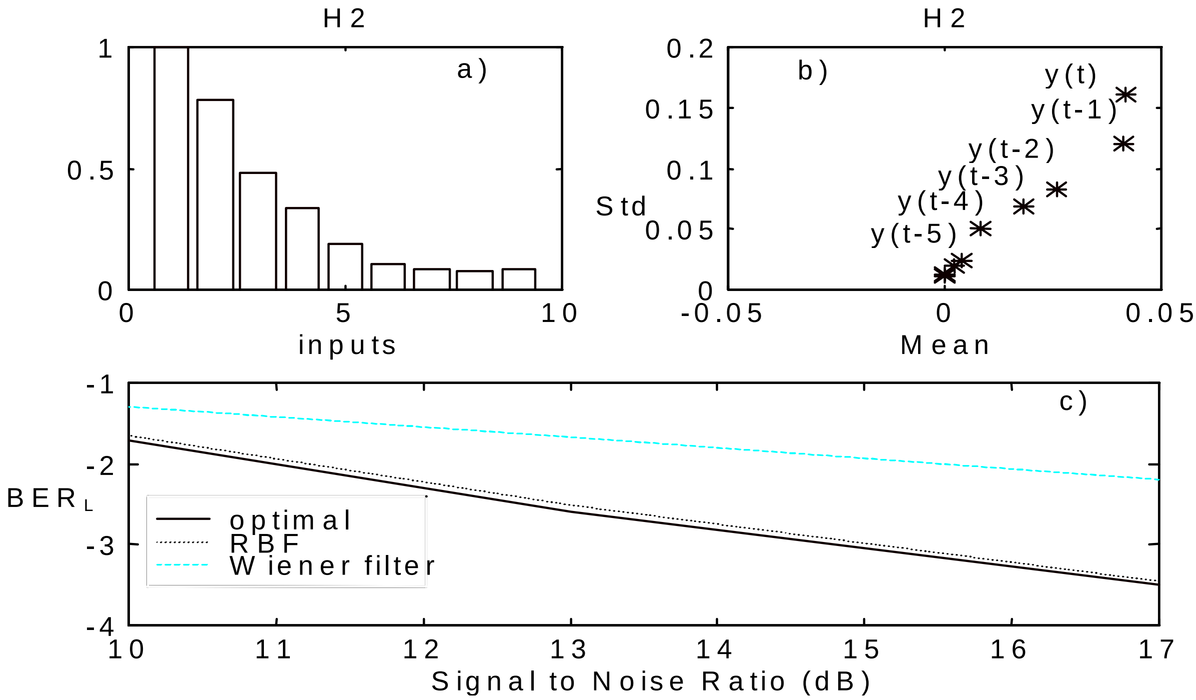

Fig. 4 – $H_2(z)$ and m=6: a) 95% normalized quantile of the absolute derivatives.
b) Mean$\left(d\bar{s}/dy\right)$ x Std$\left(d\bar{s}/dy\right)$. c) BER_L. The order of the Wiener filter equalizer is 6.