OPTIMIZATION OF A NEURAL NETWORK APPLIED TO PULSED RADAR DETECTION

Diego Andina and José L. Sanz-González
Departamento de Señales, Sistemas y Radiocomunicaciones, ETSI de Telecomunicación
Universidad Politécnica de Madrid, Ciudad Universitaria s/n 28040 Madrid, Spain,
Tel: +34 1 549 5700 (ext.380); fax +34 1 543 9652,
e-mail: andina@gc.ssr.upm.es

ABSTRACT
The purpose of this paper is to present the results of the optimization of some features of a Neural Network applied to the binary detection problem. We present how to design the structure and the training sets, and how to modify the BackPropagation algorithm to improve the results of the network for the binary detection task. The Neural Network, so designed, presents an optimal range of pulse integration and a performance very close to the theoretical limits, even under input distributions different from those used for training.

1 THE NEURAL NETWORK AS DETECTOR

The detector that we analyze is a modified envelope detector [6]. So, the input is modeled at time t as a complex value composed of

$$x_i(t) = s_i(t) + n(t), \quad i=1,2 \qquad (1)$$

where x_i is the input vector, s_i is the signal vector (s_1 corresponds to "0" and s_2 corresponds to "1") and n is the noise vector. Each component of x_i corresponds to the complex envelope of the received signal.

At the neural network output, values in (0, 1) are obtained. Then, a threshold value $T \in (0, 1)$ is chosen so that output values in (0,T) will be considered as binary output 0 and values in [T,1) will represent value 1.

The type of network used is a Multilayer Perceptron (MLP), where the typical Back-Propagation (BP) algorithm with momentum [4], that minimizes sum-of-squares is modified by the use of a more appropriate criterion function [2] purposed by El-Jaroudi and Makhoul [3], as it is explained in point 3.2.

The inputs to the network are samples of the complex envelope constituted by a sequence of M complex samples. We define the two detection hypothesis

$$H_0:\ x(kT_0) = n(kT_0) \qquad (2a)$$
$$H_1:\ x(kT_0) = S \cdot e^{j\Theta(k)} + n(kT_0) \qquad (2b)$$

where T_0 is the sample repetition period, k varies from 0 to the number of samples (M-1), S is the signal amplitude, Θ indicates an initial phase (constant in the same sequence) and $n(kT_0)$ represents an uncorrelated zero-mean Gaussian variable with variance σ^2. Each complex input is separated in its real and imaginary parts, yielding two real inputs to the net; so the number of input units must be 2M.

2 DESIGN OF THE STRUCTURE

The network has one output node, the binary output value (after thresholding), is 0 or 1. The number of inputs is equal to 2M, being M the number of samples used in each detection. The number of hidden layers and number of nodes in each layer has to be found empirically.

By adding more than one hidden layer, we have not found any performance improvements with the drawback of increasing the physical and operational complexity and increasing training and operational times, too. To choose the size of this hidden layer, we have constructed empirical curves, finding that the curves present a knee, where the most effective relationship between complexity and performance is located. After a thorough study of the values of M, N and TSNR we have found (Figures 1,2) that the TSNR parameter is critical to the performance of the net. If this parameter is properly chosen, the performance of the net can be quasi optimal (Fig.2). But this excellent performance also depends in the number pulses (M) used for each detection. If we use less than M=4 pulses for each detection (Fig. 1), the network is too small to solve the problem efficiently. If M is too high, the net is over-dimensioned, and the empirical curves separates from the optimum ones.

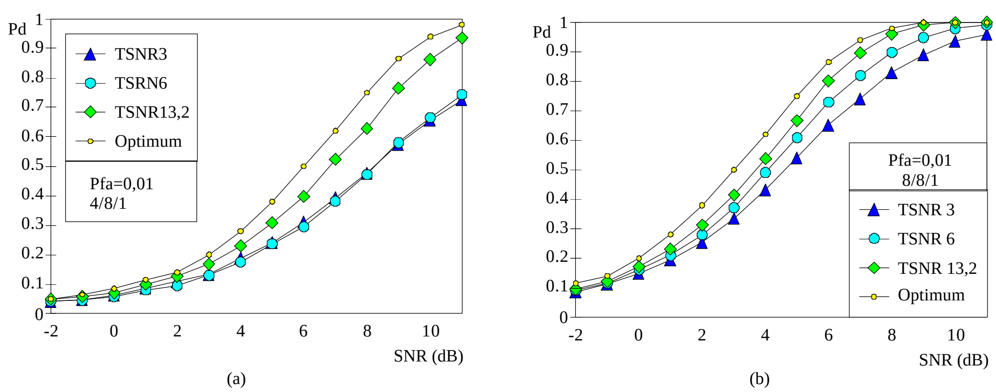

Fig. 1. Detection Probability (P_d) vs Signal-to-Noise Ratio (SNR) for a MLP of structure (a) 4/8/1 (4 inputs, 8 nodes in the hidden layer and one output) (b) 8/8/1 and different Training Signal-to-Noise Ratio (TSNR). P_{fa}= 0.01.

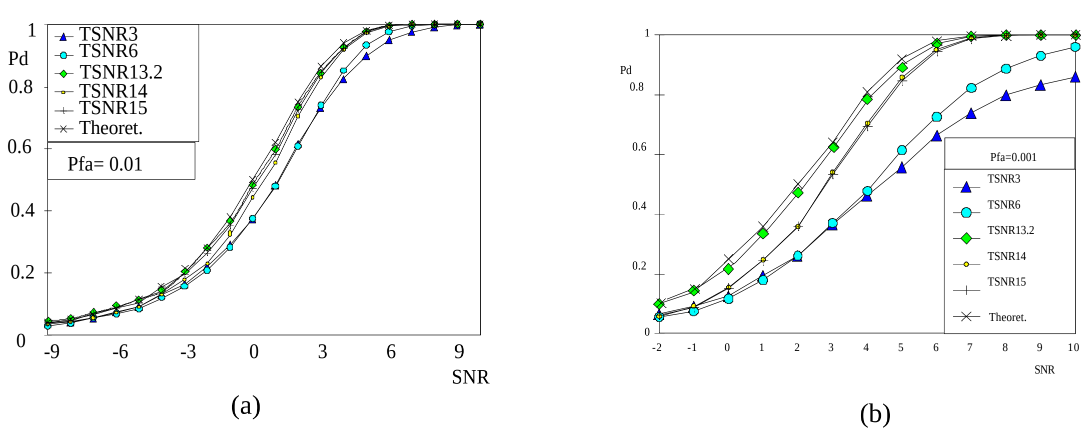

Fig. 2. Detection Probability (P_d) vs Signal-to-Noise Ratio (SNR) for a MLP of structure 16/8/1 and different Training Signal-to-Noise Ratio (TSNR). (a) P_{fa}= 0.01. (b) P_{fa}= 0.001.

3 TRAINING METHOD

3.1 Design of the training sets.

The training data has to be splitted into two sets: a training set which is used to train the network, and a test set which is used to estimate the error probability (P_e) of the neural network detector, testing its generalization capability at the same time. The training set is composed of equal number of signal-plus-noise and only-noise patterns that are presented alternatively, so that the desired output alternates between 0 and 1 at each iteration. Other choices of presenting the training pairs, as in a random sequence, are also suitable. The phase Θ is randomly varied in the interval $[0,2\pi)$ for imposing the network to generalize on the input phases during the learning procedure. The training and testing pairs have been simulated in the computer for different values of the training signal-to-noise ratio (TSNR), one of the training parameters that has showed more influence in the performance of the net. The proper relationship among signal-to-noise ratio for training (TSNR), detection probability (P_d) and false alarm probability (P_{fa}) must be found. As a general rule, the net should be trained searching the TSNR that minimizes the error probability. Then the threshold value has to be adjusted to achieve the design requirements of P_{fa}. If this is impossible, a higher training signal-to-noise ratio has to be used.

3.2 Modification of the Backpropagation algorithm.

The Backpropagation (BP) algorithm modify the output layer weights during the training as

$$w_{ij}^{(L)}(n+1) = w_{ij}^{(L)}(n) + + \eta \frac{\partial \epsilon(W)}{\partial \hat{y}_j^{(L)}} \cdot f_j^{(L)\prime} \cdot \hat{y}_i^{(L-1)} \quad \textbf{(3)}$$

where $w^{(L)}_{ij}$ is the weight connecting output from node j of the layer L-1 to the input i of the nodes in output layer L. The iteration counter is n, η is the training coefficient, $\boldsymbol{y}_j^{(L)}$ the output of the j-node of the output layer L and $f_j^{(L)'}$ the function performed by this node. $\varepsilon(W)$ is the criterion function, that depends on the weights matrix, W

For the binary detection task, it has been found that the use of the criterion function purposed by El-Jaroudi and Makhoul (JM) [3], performs better than the classical sum-of-square-error. In our case, the partial derivate of $\varepsilon(W)$, where $\varepsilon(W)$, is estimated over each input x_p can be expressed as

$$\frac{\partial \epsilon(W)}{\partial \hat{y}_j^{(L)}} = -\frac{sgn(y_j^{(L)} - \hat{y}_j^{(L)})}{1 - |y_j^{(L)} - \hat{y}_j^{(L)}|} \quad \textbf{(4)}$$

being $y_j^{(L)}$ the desired output.

4 COMPUTER RESULTS

In this point, theoretical and empirical detection curves (P_d vs SNR) and Receiver Operating Characteristics (ROC curves) are presented. For radar models as is the classical Swerling I (SWI) where S in ec.(2) has Raileigh distribution, the results are also very near to optimal, as showed in Figure 3a.

The network present also excellent performance when the network is trained with different input distribution than that one used for training (Figure 3b). A net trained by the simple Marcum mode is evaluated over the SWI input distribution. The result is that the network is not only capable of generalize an to obtain good results over the more complicated input mode SWI, but also that its results are superior to those of the network trained over the SWI distribution.

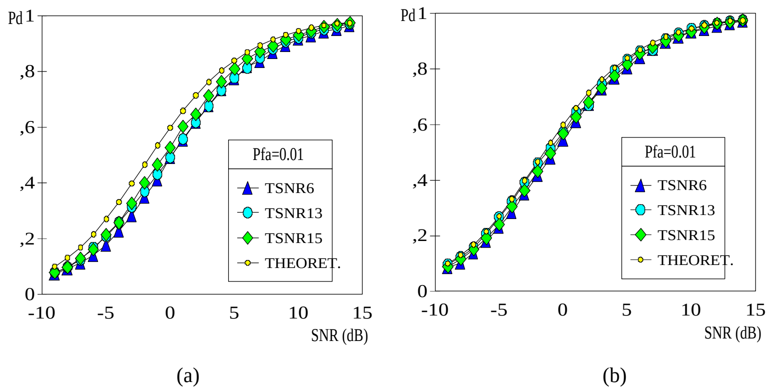

Fig. 3. Detection Probability (P_d) vs Signal-to-Noise Ratio (SNR) for a MLP of structure 16/8/1 (16 inputs, 8 nodes in the hidden layer and one output) and different Training Signal-to-Noise Ratio (TSNR). P_{fa}= 0.01. (a) Radar Swerling I Model for training and testing. (b) Net trained with Marcum model and tested over Swerling I input samples.

5 CONCLUSIONS

(a) A Neural Network applied to binary detection is optimized in several ways. The purposed structure is a three layer MLP, using the adequate criterion function and optimizing the relationship among signal-to-noise ratio for training (TSNR), detection probability (P_d) and false alarm probability (P_{fa})

(b) The net should be trained searching the TSNR that minimizes the error probability. Then the threshold value has to be adjusted to achieve the design requirements of P_{fa}. If this is impossible, a higher training signal-to-noise ratio has to be used.

(c) In the purposed detector, the number of inputs depends on the number of pulses to be integrated. But it has been found that there is an optimal range on this number of pulses. The lower limit of this range will correspond to a too simple network. The upper limit will correspond to an over-dimensioned network.

(d) At least for the input model presented, and following the upper conclusions, the empirical results of the purposed detector reach the theoretical limits, even when the input distribution is different than that one used to design the Neural Network.

6 REFERENCES

[1] **D. Andina, J.L. Sanz-González**, "On the problem of Binary Detection with Neural Networks", *Proc. of 38 Midwest Symposium on Circuits and Systems*, Rio de Janeiro, Brazil, August 1995.

[2] **D. Andina, J.L. Sanz-González and J.A. Jiménez-Pajares**, "A Comparison of Criterion Functions for a Neural Network Applied to Binary Detection", Accepted for publication in *Proc. of Int. Conf. on Neural Networks, ICNN*, Perth, Australia,pp. 329-333, November 1995.

[3] **A. El-Jaroudi and J. Makhoul**, "A New Error Criterion For Posterior Probability Estimation With Neural Nets", *Proc. of Int. Joint. Conf. on Neural Networks, IJCNN*, Vol I, No. 5, pp. 185-192, 1990.

[4] **D.R. Hush and B.G. Horne**, "Progress in Supervised Neural Networks. What´s new since Lippmann ?", *IEEE Signal Processing Magazine*, pp. 8-51, January 1993.

[5] **J.L. Marcum**, "A Statistical Theory of Target Detection by Pulsed Radar". *IRE Trans. on Information Theory*, Vol. IT-6, Nº 2, pp. 59-144. April 1960, .

[6] **W.L. Root**, "An Introduction to the Theory of the Detection of Signals in Noise". *Proc. of the IEEE*, vol 58, pp. 610-622, May 1970.