An efficient channel equalizer using artificial neural networks

Article *in* Neural Network World · January 2006

Amiya Kumar Tripathy
Don Bosco Institute of Technology

An Efficient Channel Equalizer Using Artificial Neural Networks

K. T. Raghavendra
Don Bosco Institute of Technology, Mumbai, India
raghavendrakt@yahoo.com

Amiya K. Tripathy
Don Bosco Institute of Technology, Mumbai, India
triapthy.a@gmail.com

ABSTRACT

When digital signals are transmitted through frequency selective communication channels, one of the problems that arise is inter-symbol interference (ISI). To compensate corruptions caused by ISI and to find the original information being transmitted, equalization process is performed at the receiver. Since communication channels are time varying and random in nature, adaptive equalizers must be used to learn and subsequently track the time varying characteristics of the channel. Traditional equalizers are based on finding the inverse of the channel and compensating the channel's influence using inverse filter technique. There exists no equalizer for non-invertible channels. Artificial Neural Networks (ANN) can be applied to this for achieving better performance than conventional methods. We have proposed a model of neural equalizer using MLP (multi layer perceptron), which reduces the mean square error to minimum and eliminates the effects of ISI. Empirically we have found this neural equalizer is more efficient than conventional adaptive equalizers.

Key Words: Adaptive Channel Equalizer, Artificial Neural Networks, Multi Layer Neural Network and Inter Symbol Interference.

1. Introduction

The task of a receiver is to retrieve the information send by the transmitter. In order to do that it has to determine what has been send. To accomplish this task, it tries to extract from the receive signal the parameters related to the transmitted information. However there are some parameters that are not the actual part of the transmit data, but are included by the medium which is used for transmission of the signal. One of such parameters is the channel. The transmit signal passes through the channel before reaching the receiver, or in other words the transmit signal convolves with the channel.

When transmitting data, e.g. fax, video streams, or web contents, over the copper cable, the transmitted symbols suffer from (ISI), which is nothing but the convolution of transmit signal with the channel [3]. The responses of consecutively transmitted data symbols influence each other at the receiver due the dispersive nature of the channel. ISI is introduced by the delay spread inherent in frequency-selective channels. The response of a transmitted symbol overlaps with the response of the next symbol in time domain, which makes the reception more complicated. Specifically, if the time duration of a transmitted pulse is much larger than the multi-path delay spread, all multi-path components arrive at the receiver approximately on top of each other; consequently, the received narrowband pulse experiences a very small amount of time spread and interferes negligibly with other pulses [1]. However, if the time duration of the transmitted pulse is much smaller than the multi-path delay spread, then each of the multi-path components can be resolved in time at the receiver. Therefore, interference with previously and subsequently transmitted pulses (ISI) occurs which has been found to produce irreducible error floors in most digital modulation techniques.

The principle in time domain has shown in Figure 1. The positive-valued response overlaps with the negative-valued response as shown.

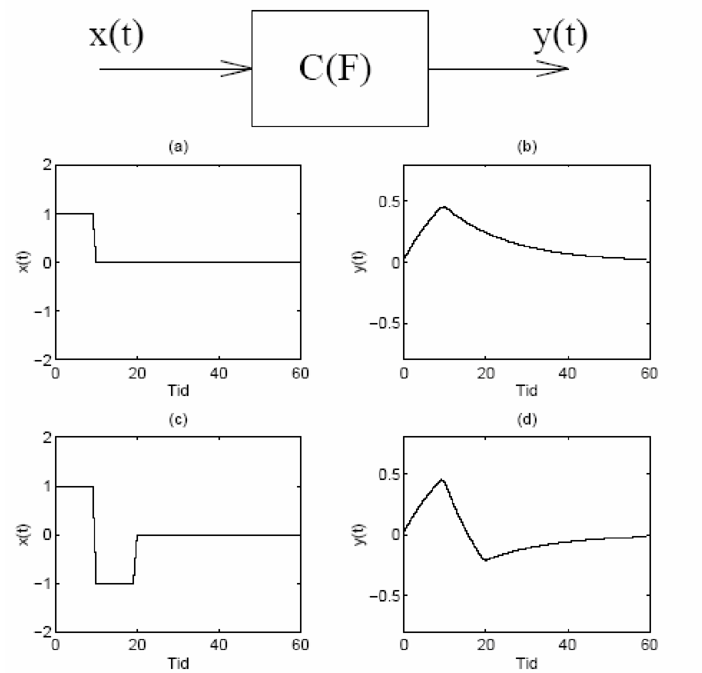

Figure 1. Illustration of ISI. (a) and (c) are input signals, (b) and (d) are the corresponding output signal of the channel C (F).

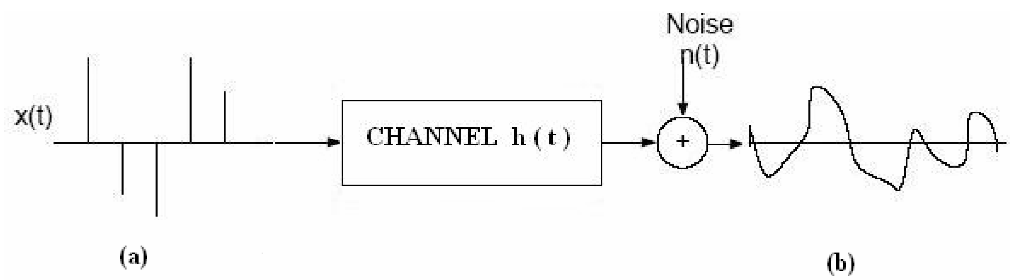

Figure 2. Effect of ISI. (a) Transmitted signal (b) Received signal.

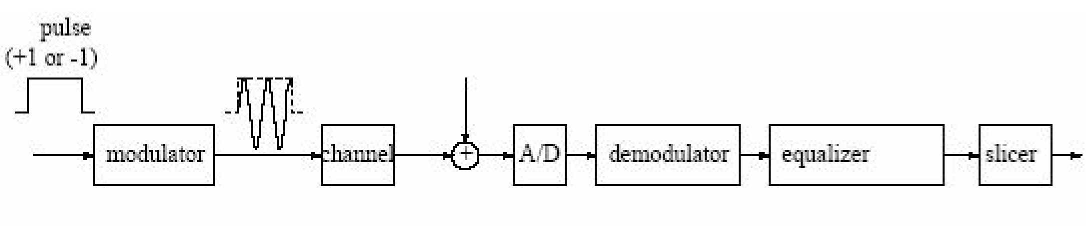

Figure3. Model of data transmission system.

Where x(t) is the input signal, y(t) is the out put signal and C(F) is the transfer function of the system. Figure 2 shows the effect of Inter symbol Interference (ISI) in a communication system. It is desirable to find a mathematical model of the channel. A copper cable, e.g., can roughly be modeled as a band pass filter with a pass band of 300 − 3300Hz. An adequate mathematical model of the channel constitutes the basis for compensating the channel's influence, which is referred to as channel equalization. In case the inverse of the channel exists, the channel is said to be invertible. This inverse of the channel is used in channel equalization. Figure 3 shows a more realistic model of a data transmission system, which consists of a modulator, a channel, and a demodulator.

Equalization refers to any signal processing techniques used at the receiver to combat ISI in depressive channels in the presence of additive noise. Although the main purpose of equalization is to alleviate the effects of ISI, tradeoffs must be made so that noise enhancement does not result. In order for an equalizer to mitigate ISI, an estimate of the channel response is needed at the receiver. The properties of, e.g. a radio channel may depend on the atmospheric conditions, possibly existing reflecting objects, and the movement of both transmitter and receiver. A copper cable appears to be a much more static channel; however, the actual transfer function of a randomly chosen pair may differ substantially from the theoretical model since several parameters as well as their influence on the cable's behavior are unknown. The temperature of the cable, the age of the cable, and the exact arrangement of the cable with respect to ground as well as other cables and conductors are examples for such parameters. Hence, a priori an accurate channel model is not available. Once the channel properties have been determined exactly, they may change with time. This is obvious for the radio channel; however, even the properties of the copper cable may change, for example due to a change of the cable's temperature caused by the weather. Moreover wireless channels are random in nature. Thus adaptive channel equalization, i.e. equalization that can adapt to different channel conditions, learn themselves and subsequently track the time varying characteristics of the channel, is required. This adaptive channel equalizer is nothing but an adaptive filter, which has adjustable filter coefficients and these coefficients, adapts new values whenever there is a change in the output according to the algorithm employed.

In the rest of this paper, in section 2, we discuss the background of Adaptive channel equalization and some drawbacks of the present techniques. In section 3, we present our proposed model of an efficient adaptive channel equalizer using artificial neural networks. In section 4, we have shown the performance evaluation of our proposed model and its discussion. Section 5 is the conclusion.

2. Background

2.1 Principle of Adaptive Channel Equalizer

The transmitted signal passes through the channel before reaching the receiver, or in other words the transmit signal convolves with the channel. This convolution bring distortion in the transmit signal called the inter-symbol interference. The effect of the channel can be nullified by a de-convolution operation. A simple de-convolution can be realized in Fourier domain [2, 4, 15]. But, the problem here is that neither the transmit signal nor the channel are known. Therefore, a training signal is required which is known at the receiver. But, the Fourier domain approach is for deterministic signals and deterministic systems, while the input signal here is stochastic in nature plus the channel is time varying, and there is also channel noise. The extraction of channel parameters from the retrieved signal can therefore be done using Wiener solution and i.e.

$$W_{opt} = R^{-1} p \ldots\ldots (1)$$

Where, R is the autocorrelation matrix of the received data, W_{opt} is the optimum filter coefficient and p is the cross correlation matrix. We assumed channel as simple FIR filter. Here, the matrix inversion is difficult to accomplish. Method of gradient estimate is chosen to find the weights in a number of small steps rather than one step. Thus, the update equation turns out to be

$$W_{k+1} = W_k + \mu \nabla E(e_k^2) \ldots\ldots (2)$$

The gradient estimate chosen for this assignment is the stochastic gradient estimate, which reduces the expression to a simpler form

$$W_{k+1} = W_k + \mu e_k X_k \ldots\ldots (3)$$

Here, e_k is the error signal, X_k is the received signal vector W_k is the kth filter coefficient.

Before, going into implementation details it is better to probe a bit further on e_k and X_k. The figure 4 shows the error signal and the received signal. For implementation point of view it is important to mention here that the error signal e_k is a single value in equation (3), while the received signal vector X_k is not a single value but a vector that is equal in length to the filter taps N. The vector contains the current value X(k) and N-1 previous values. The kth output of the adaptive filter is actually dot product of the receive signal vector X_k and the adaptive filter taps W_k at that moment

$$y(k) = W_k T X_k \ldots\ldots (4)$$

Therefore the error signal is the difference between the output of the adaptive filter and the training signal.

$$e(k) = d(k) - y(k) \ldots\ldots (5)$$

This e(k) is written as e_k in the above mentioned equations [8,10, 13, 14, 26].

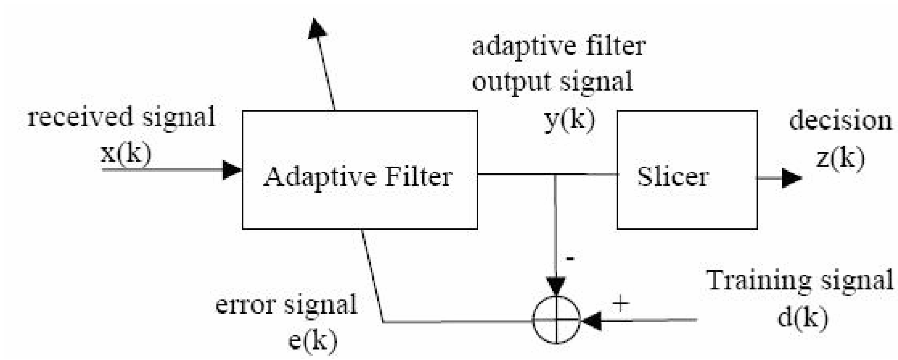

Figure 4. Model of Adaptive Channel Equalizer

2.2 Linear and Nonlinear Equalization.

Existing adaptive equalization techniques are divided into two general categories: linear and nonlinear equalization. The difference between linear and nonlinear equalizers is determined by whether the reconstructed symbol at the receiver output is employed in a feed back path to adapt the equalizer for subsequent outputs: linear equalizers do not utilize a feed back path while nonlinear equalizers do. In both linear and nonlinear equalization techniques, however, the channel model is generally assumed to be linear. Furthermore, equalizers are characterized by their architectures and training algorithm. Linear equalizers generally employ linear filters with transversal or lattice structures and adaptation algorithms such as recursive least squares (RLS), least mean squares (LMS), fast RLS, square-root RLS, gradient RLS, etc. However, linear equalizers do not perform well on channels with deep spectral nulls since the equalizer places a large gain at these frequencies and consequently significantly enhances the additive noise [7]. Noise enhancement can be overcome by using nonlinear equalization techniques such as decision feedback equalization (DFE) or maximum likelihood sequence estimation (MLSE).

2.3 Some Drawbacks in Existing Systems

If the channel suffers from significant nonlinear distortions, the aforementioned equalization techniques exhibit poor performance, and equalization techniques that better combat nonlinear channels are desirable. The adaptive equalization problem is typically viewed as an inverse filter problem. Traditional equalizers are based on finding the inverse of the channel and compensating the channel's influence using inverse filter technique. These equalizers are designed to approximately track and invert time-varying channel distortions by adjusting filter coefficients while maintaining a prescribed signal to noise ratio (SNR). In many practical cases the channel cannot be inverted which may have various reasons [21, 23]. Hence there exists no equalizer for these non-invertible channels. This is the major disadvantage with normal adaptive channel equalizers. The output of the equalizer is fed into a decision device, which attempts to estimate the transmitted symbol. This configuration of an inverse filter equalizer followed by a decision device results in the partitioning of the output signal space by linear decision boundaries between different symbols. When significant noise is added to the transmitted signal, linear boundaries are not optimal.

3. A Proposed Model

Equalization can be considered as a geometric classification problem rather than an inverse filter problem; however, the main objective becomes the separation of the received symbols in the output signal space whose optimal decision region boundaries are generally highly nonlinear. The idea here is to classify the received signal vectors by partitioning the signal space into some decision regions. With this viewpoint to equalization, complete channel inversion is unnecessary, and the problem is tackled using classification techniques.

Artificial Neural Networks (ANN) can be applied to this field for achieving better performance than classical methods in some aspects. Since Artificial Neural Networks are well known for their ability of performing classification tasks by forming complex nonlinear decision boundaries, neural equalizers have been recently receiving considerable attention in order to increase receiver robustness. Neural equalizers have the potential for significant performance improvements especially in severely distorted, nonlinear channels [15,16,17,19]. Artificial Neural Networks are parallel distributed processing systems in which many simple interconnected elements (neurons) simultaneously process information, adapt and learn from past patterns [2, 3, 8, 9]. Although only capable of performing simple operations themselves, when organized into layers, neurons are collectively capable of performing highly sophisticated operations. Attractive properties of ANN that are relevant to the equalization problem at hand include massive parallelism, adaptive processing, self-organization, universal approximation, and most importantly, the capability of tackling highly nonlinear problems.

When using neural equalizer the problem of channel inversion does not exist and hence they can be used with all channels. Due to this major advantage Neural Equalizer finds importance in all telecommunication systems [11, 12]. Our proposed model of adaptive channel equalizer consists of multi layer perception network, which reduces the Mean Square Error (MSE) to minimum, and it can handle highly non-linear communication channels.

The multilayer perceptron (MLP) is a popular neural network architecture that is frequently used for highly nonlinear classification problems [6, 7, 14]. The basic element of the MLP is the neuron, and a network is formed by arranging these neurons into layers as depicted in Figure 2(a). Each neuron, or node, in the network is composed of a linear combiner and activation function as portrayed in Figure 2(b). Each neuron receives the inputs $x_1, x_2, \ldots, x_N$ which are multiplied by the synaptic weights $w_1, w_2, \ldots, w_N$, respectively. The output of the linear combiner is the weighted sum of the inputs plus a threshold or bias term, b. Finally; the activation function gives the output of each neuron as

$$x = f(y) \text{ where } \quad y = \sum_{i=1}^{N} w_i x_i + b$$

Where x_i is the i^{th} input, w_i the corresponding weight, and b the bias. Note that the weights, which attenuate the inputs, and the bias term, specify classification boundaries within the network. The activation function, $f(\cdot)$, may be a linear function, but generally a nonlinear function (i.e. sigmoidal, logistic, hyperbolic tan) is desirable since it improves learning for nonlinear channels.

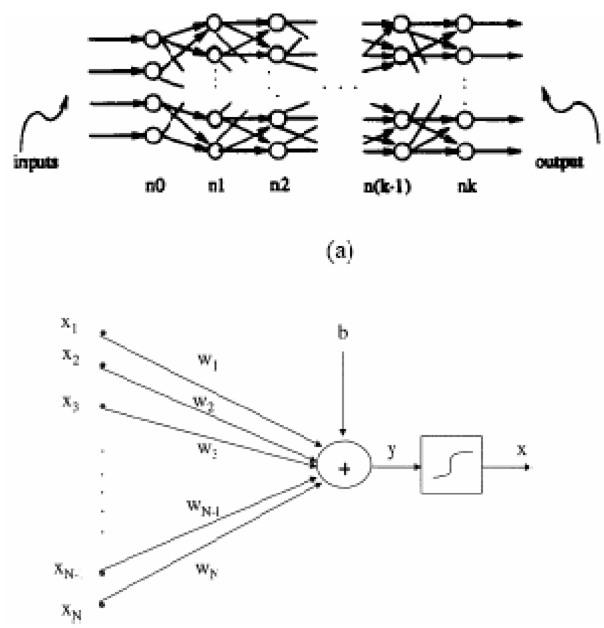

Figure 5. (a) Multilayer Perceptron rchitecture (b) the neuron

The inputs to the MLP network (i.e. several delayed versions of the channel output) are the inputs to the first layer of neurons while the overall network outputs are the outputs of the output layer. Neurons are cascaded in layers so that the output of one layer is the input to the next sequential layer. The layers between the input and output layers are known as the hidden layers which are the decision layers that are collectively capable of performing complex nonlinear mappings between the input and output layers. More specifically, the hidden layers exploit the (nonlinear) activation functions to create the complex nonlinear decision regions of the signal space. Generally, in a MLP network, all neurons in a layer are fully connected to neurons in adjacent layers, but not connected within a layer. Moreover, there are no connections bridging layers. Combinational operations in the hidden layers generate the overall output of the network. As represented in Figure 2(a), the MLP architecture is described by the sequence of integers n_o, n_1,…, n_k and is called MLP(n_o, n_1,…, n_k) where n denotes the number of neurons in each successive layer from input to output. Consequently, it can be seen that the MLP performs the nonlinear mapping: $R^{n0} \rightarrow R^{nk}$.

The "information" of the channel is stored in the weights, or coefficients, of the neurons, and these weights must be periodically updated to track the time varying channel. The weights are generally updated by using the back-propagation (BP) training algorithm, which is a gradient descent optimization procedure. Since the BP algorithm is a supervised learning algorithm, a set of actual output/desired output pairs train the network to implement the desired mapping. The BP algorithm is iterative and adjusts the weights so as to minimize any differentiable cost function such as the mean square error (MSE). Since the BP algorithm performs a gradient descent to reach a minimum, the search direction is given by the steepest gradient at the current search point, and the weights are adjusted by an amount proportional to local gradients. For complex valued signals, the complex version of the back-propagation algorithm (CBP) is employed.

4. Evaluations and Discussion

Figure 6 shows the Impulse response of the channel. A realistic nonlinear channel model is utilized with the nonlinearity arising from the use of efficient nonlinear power amplifiers and the16-QAM modulation scheme is used. Figure 7 shows the response of the equalizer.

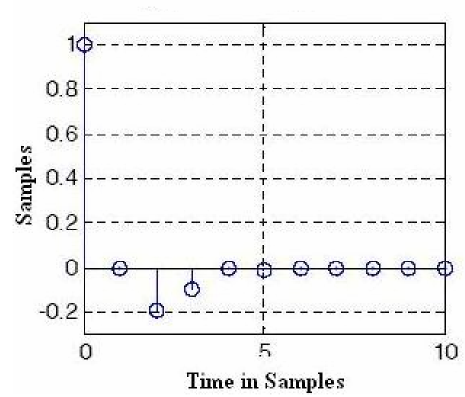

Figure 6. Response of the channel

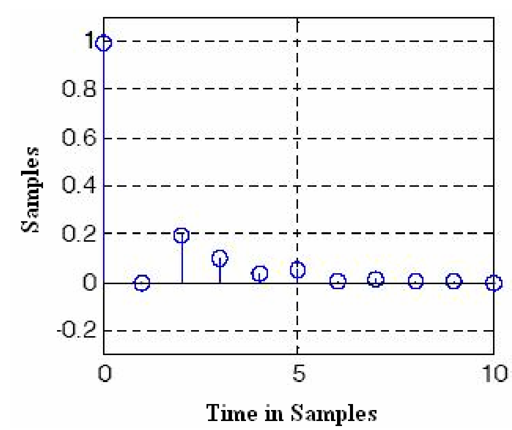

Figure 7. Response of the Equalizer

Figure 8 shows the overall response (convoluted response) of communication system with equalizer compensating channel impairments. In case of ideal response of the channel multiplied with that of equalizer should be unity for all the frequencies, this is always not possible as the number of taps that can exactly represent the channel might be unknown, there is channel noise and there can be error in estimating the group delay. All the above mentioned parameters would lead to erroneous estimation of the inverse channel, and depending on the magnitude of the error the product of the channel response and the equalizer response would deviate from the ideal of all unity for all frequencies.

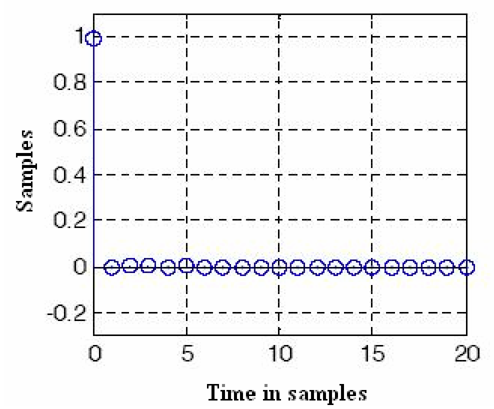

Figure 8. Overall response of communication system.

We have performed this experiment by taking randomly chosen data (input signal) and feeding it to the channel, which is modeled as simple FIR filter. The output of FIR filter is given to the LMS equalizer. Figure 9 shows the MSE, which lies in the range 10^{-1} to 10^{-2}.

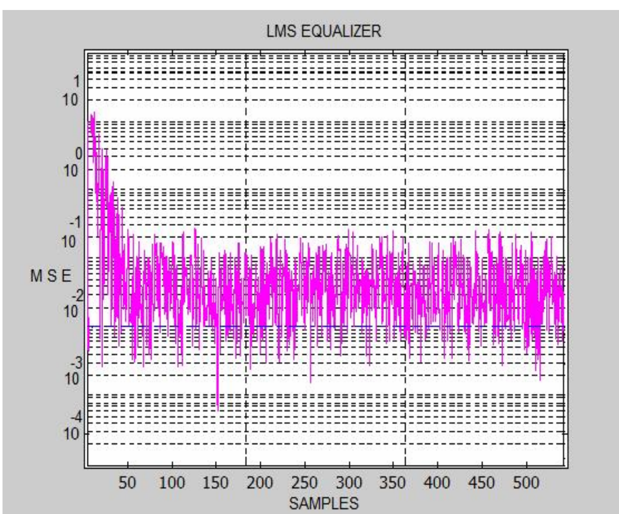

Figure 9. The error (MSE) between the actual desired signal and the output of LMS equalizer.

Figure 10 shows MSE of neural (MLP) equalizer. Simulation results show that neural equalizer using MLP network is more efficient than conventional LMS equalizer. Neural equalizer performs well in distorted and nonlinear channels. Simulation results also show that performance of neural (MLP) equalizer is better than LMS equalizer with respect to error minimization i.e. MSE lies in the range of 10^{-2} to 10^{-3}.

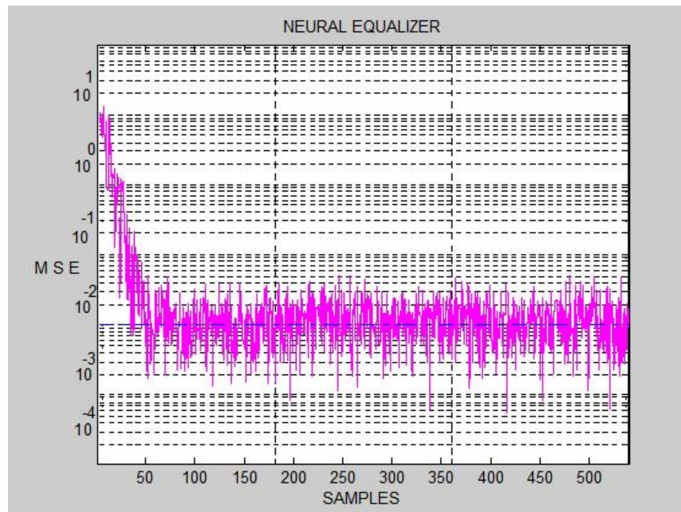

Figure 10. The error (MSE) between the actual desired signal and the output of neural (MLP) equalizer.

5. Conclusion

Based on the improved performance of neural network with its successful applications to the nonlinear channel equalization problem, we have introduced an MLP-based equalization Model. This model is applied to a nonlinear channel. We have carried out simulations with the 16-QAM under various conditions pointed out that this deformation in the signal space could be successfully reversed to achieve practically acceptable bit error rates. Also comparison of the MLP equalizer with the MSE equalizer demonstrated the fact that the error produced is minimum in case of MLP. A few noticeable things of this equalizer are its computational simplicity, due to the small size of MLPs that can achieve good performance and efficient extraction of information from a small number of training samples.

6. Acknowledgement

This work is fully supported by Don Bosco Institute of Technology, Mumbai, India.

References

[1] Andre Tkacenko, and P. P. Vaidyanathan, "A Low-Complexity Eigen filter Design Method for Channel Shortening Equalizers for DMT Systems", IEEE Transaction on Communication, July 2003, Vol. 51, No.7.

[2] Albu, A. Mateescu, J. C. M. Mota, B. Dorizzi, "Adaptive Channel Equalization Using Neural Network," in Proc. ITS '98—SBT/IEEE Int. Telecommunications Symposium, 1998, vol. 2, pp. 438-441.

[3] A. Al-Mashouq, I. S. Reed, "The Use of Neural Nets to Combine Equalization with Decoding for Severe Inter Symbol Interference Channels," IEEE Trans. on Neural Networks, 1994, vol. 5, issue 6, pp. 982-988.

[4] Cha and S. A. Kassam, "Channel Equalization Using Adaptive Complex Radial Basis Function Networks," IEEE Journal on Selected Areas in Communications, 1995, vol. 13, issue 1, pp. 122-131.

[5] Deng, S. Narasimhan, P. Saratchandran, "Communication Channel Equalization Using Complex-Valued Minimal Radial Basis Function Neural Networks," IEEE Trans. on Neural Networks, 2002, vol. 13, issue 3, pp. 687-696.

[6] Erdogmus, D. Rende, J. C. Principe, T. F. Wong, "Nonlinear Channel Equalization Using Multilayer Perceptrons with Information-Theoretic Criterion," in Proc. of 2001 IEEE Signal Processing Society Workshop, 2001, pp. 443-451.

[7] F. N. Cowan, "Nonlinear Adaptive Equalisation (Multilayer Perceptron)," Sixth Int. Conf. on Digital Processing of Signals in Communications, 1991, pp. 1-5.

[8] Gan, P. Saratchandran, N. Sundararajan, K. R. Subramanian, "A Complex Valued Radial Basis Function Network for Equalization of Fast Time Varying Channels," IEEE Trans. on Neural Networks, 1999, vol. 10, issue 4, pp. 958-960.

[9] Henrique and G. Coelho, "Adaptive channel equalization using EKF-CRTRL neural networks", Proceedings of the 2002 International Joint Conference on Neural Networks, IJCNN '02, 2002, Vol.2, pp. 1195-1199.

[10] J. Gibson, S. Siu, S. Chen, C. F. N. Cowan, P. M. Grant, "The Application of Nonlinear Architectures to Adaptive Channel Equalization," in Proc SUPERCOMM/ICC '90—IEEE Int. Conf. on Communications, 1990, vol. 2, pp. 649-653.

[11] J. C. Patra, R. N. Pal, R. Baliarsingh, G. Panda, "Nonlinear Channel Equalization for QAM Signal Constellation Using Artificial Neural Networks," IEEE Trans. on Systems, Man, & Cybernetics, Part B, 1999, vol. 29, issue 2, pp. 262-271.

[12] Kechriotis, E. Zervas, E. S. Manolakos, "Using Recurrent Neural Networks for Adaptive Communication Channel Equalization," IEEE Trans on Neural Networks, 1994, vol. 5, issue 2, pp. 267-278.

[13] Krstajic, Z. Uskokovic, and L. Stankovic; "Adaptive channel equalizer with new VSS LMS algorithm", Proceedings. Seventh International Symposium on Signal Processing and Its Applications; 2003, vol. 2, pp. 567 - 570.

[14] Kantsila, T. Haverinen, M. Lehtokangas, J. Saarinen, “On Equalization with Complex MLP Network in the GSM Environment,” in Proc. Joint 9th IFSA World Congress and 20th NAFIPS Int. Conf, 2001, vol. 5, pp. 2687-2692.

[15] Lee, C. D. Beach, N. Tepedelenlioglu, “Channel Equalization Using Radial Basis Function Network,” IEEE Int. Conf. on Neural Networks, 1996, vol. 4, pp. 1924-1928.

[16] Peng, C. L. Nikias, J. G. Proakis, “Adaptive Equalization with Neural Networks: New Multi-Layer Perceptron Structures and Their Evaluation,” in Proc. ICASSP ’92 - IEEE Int. Conf. Acoustics, Speech, & Signal Processing, 1992, vol. 2, pp. 301-304.

[17] Peng, C. L. Nikias, J. G. Proakis, “Adaptive Equalization for PAM and QAM Signals with Neural Networks,” in Proc. of 25th Asilomar Conf. on Signals, Systems & Computers, 1991, vol. 1, pp. 496-500.

[18] Power, F. Sweeney, C. F. N. Cowan, “Non-Linear MLP Channel Equalization,” IEE Colloquium on Statistical Signal Processing (Ref. No. 1999/002), 1999, pp. 3/1-3/6.

[19] Parisi, E. D. Di Claudio, G. Orlandi, B. D. Rao, “Fast Adaptive Digital Equalization by Recurrent Neural Networks,” IEEE Trans on Signal Processing, 1997, vol. 45, issue 11, pp. 2731-2739.

[20] Siu, G. J. Gibson, C. F. N. Cowan, “Decision Feedback Equalization Using Neural Network Structures,” First IEE Int. Conf. on Artificial Neural Networks (Conf. Publ. No. 313), 1989, pp. 125-128.

[21] S. Chen, B. Mulgrew, P. M. Grant, “A Clustering Technique for Digital Communications Channel Equalization Using Radial Basis Function Networks,” IEEE Trans. on Neural Networks, vol. 4, issue 4, pp. 570-590, 1993.

[22] W. R. Kirkland, D. P. Taylor, “On the Application of Feed Forward Neural Networks to Channel Equalization,” in Proc IJCNN ’92—Int. Joint Conf. on Neural Networks, 1992vol. 2, pp. 919-924.

[23] Wan De Weng, Rui Chang Lin and Chung Ta Hseh, “The Design of an SCFNN Based Nonlinear Channel Equalizer” Journal of Information Science and Engineering 21, 2005, pp. 695-709.

[24] W. K. Lo, H. M. Hafez, “Neural Network Channel Equalization,” in Proc IJCNN ’92—Int. Joint Conf. on Neural Networks, 1992, vol. 2, pp. 981-986.

[25] Amita Pugalia, ”Adaptive Channel Equalization Using Neural Networks”, Class Projects on Wireless Communication, fall 2002, Stanford University.

[26] http://www.dspexperts.com