Optical Performance Monitoring by Use of Artificial Neural Networks Trained with Parameters Derived from Delay-Tap Asynchronous Sampling

Jeffrey A. Jargon[1], Xiaoxia Wu[2], and Alan E. Willner[2]
[1]*National Institute of Standards and Technology, Boulder, CO 80305 USA*
[2]*University of Southern California, Los Angeles, CA 90089 USA*
jargon@boulder.nist.gov

Abstract: We demonstrate a technique for optical performance monitoring by simultaneously identifying optical signal-to-noise ratio (OSNR), chromatic dispersion (CD), and polarization-mode dispersion (PMD) using artificial neural networks trained with parameters derived from delay-tap asynchronous sampling.

OCIS codes: (060.2330) Fiber optics communications; (100.4996) Pattern recognition, neural networks.

1. Introduction

As optical fiber transmission systems become more transparent and reconfigurable, optical performance monitoring (OPM) is essential for ensuring high quality of service [1]. Crucial impairments in optical networks include optical signal-to-noise ratio (OSNR), chromatic dispersion (CD), and polarization-mode dispersion (PMD).

Several techniques have been proposed for monitoring optical performance by use of off-line digital signal processing of received data signals [2-10]. Three of these methods [2-4] utilize amplitude histograms or power distributions to estimate bit error rate (BER); three [5-7] employ delay-tap plots to distinguish among impairments; and two [8-9] use parameters derived from eye diagrams for the same purpose.

Only two of them, however, have been shown to concurrently quantify at least three different impairments [6, 9]. Of these monitoring techniques, one [6] exploits pattern recognition for estimating simultaneous impairments using asynchronous, delay-tap sampling, and the other [9] uses artificial neural networks (ANNs) trained with parameters derived from synchronously sampled eye diagrams. In asynchronous sampling, the signal of interest is sampled without regard to an instant relative to a decision time, and thus clock recovery is not necessary. Synchronous sampling, however, necessitates a standard receiver with clock recovery, but can easily be used to generate eye diagrams from which numerous performance parameters may be derived.

Since ANNs have been shown to be a powerful modeling tool for identifying simultaneous impairments derived from eye-diagram parameters, we explore their use for the same purpose using parameters derived from delay-tap asynchronous sampling In the following sections, we briefly overview ANNs, and provide examples of our proposed method with a simulated optical channel operating at 10 Gbps and using non-return-to-zero, on-off keying (NRZ-OOK).

2. Artificial Neural Networks

Artificial neural networks (ANNs) are neuroscience-inspired computational tools that are trained by use of input-output data to generate a desired mapping from an input stimulus to the targeted output [10-11]. ANNs consist of multiple layers of processing elements called neurons. Each neuron is linked to other neurons in neighboring layers by varying coefficients that represent the strengths of these connections. ANNs learn relationships among sets of input-output data that are characteristic of the device or system under consideration. After the input vectors are presented to the input neurons and output vectors are computed, the ANN outputs are compared to the desired outputs, and errors are calculated. Error derivatives are then calculated and summed for each weight until all of the training sets have been presented to the network. The error derivatives are used to update the weights for the neurons, and training continues until the errors drop below prescribed values.

The ANN architecture used in this work is a feed-forward, three-layer perceptron structure (MLP3) consisting of an input layer, a hidden layer, and an output layer, as shown in Figure 1. The hidden layer allows complex models of input-output relationships. The mapping of these relationships is given by $\mathbf{Y} = g[\mathbf{W}_2 \bullet g(\mathbf{W}_1 \bullet \mathbf{X})]$, where $\mathbf{X}$ is the

This work is supported by the U.S. Department of Commerce and the DARPA CORONET program, and is not subject to U.S. Copyright.

input vector, **Y** is the output vector, and $\mathbf{W}_1$ and $\mathbf{W}_2$ are respectively the weight matrices between the input and hidden layers and between the hidden and output layers. The function $g(u)$ is a nonlinear sigmoidal activation function given by $g(u)=1/[1+\exp(-u)]$, where u is the input to a hidden neuron. According to [12], an MLP3 with one hidden sigmoidal layer is able to model almost any physical function accurately, provided that a sufficient number of hidden neurons are available.

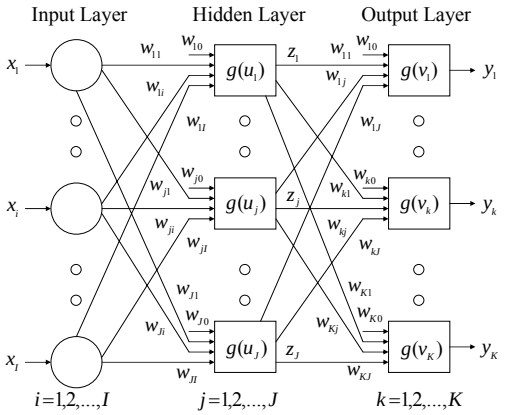

Fig. 1. Artificial neural network architecture.

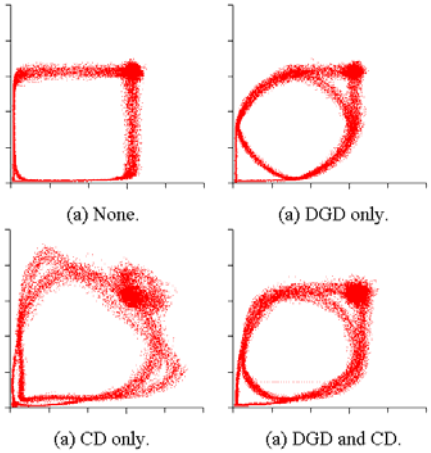

Fig. 2. $B/2$ (one-half bit-period) delay-tap plots with various impairments.

3. Methodology

Recently, Dods and Anderson introduced an asynchronous sampling technique based on a two-tap delay line, where each sample point is comprised of two measurements separated by a specific period corresponding to the length of the delay [5]. By creating a scatter plot of the measured pairs, they observed that delay lengths of less than half of the bit period ($B/2$) represented the power evolution within each bit. Furthermore, they showed that plots making use of delays of $B/2$ highlight distortion effects. Figure 2 illustrates simulated B/2 delay-tap scatter plots for a 10 Gbps NRZ-OOK signal at a few select combinations of CD and PMD for a given value of OSNR. Visually, it is obvious that these impairments produce distinct features.

To quantify the distinct features, we need to derive parameters that can be calculated from the delay-tap plots. Whereas eye diagrams give rise to widely used parameters such as Q-factor, closure, jitter, and crossing amplitude, there are no such parameters available for delay-tap plots. Thus, we propose new parameters that will help us to capture the behavior of such plots. One possibility is to divide the plots into four quadrants, Q1-Q4. The data pairs are divided into the quadrants as follows: $(x_i, y_i) \in \text{Q1}$ if $\{0 \le x_i \le \text{Max}(x)/2 \text{ and } 0 \le y_i \le \text{Max}(y)/2\}$; $(x_i, y_i) \in \text{Q2}$ if $\{0 \le x_i \le \text{Max}(x)/2 \text{ and } \text{Max}(y)/2 < y_i \le \text{Max}(y)\}$; $(x_i, y_i) \in \text{Q3}$ if $\{\text{Max}(x)/2 < x_i \le \text{Max}(x) \text{ and } \text{Max}(y)/2 < y_i \le \text{Max}(y)\}$; and quadrant 4 is not used in this case because it contains data that are the mirror image of quadrant 2. Figure 3 illustrates this concept.

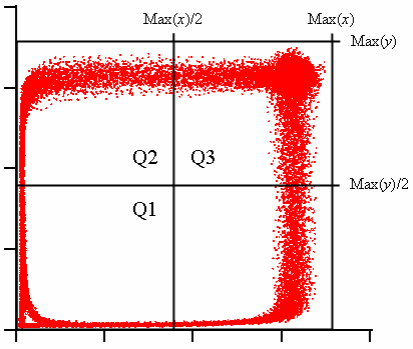

Fig. 3. Dividing the delay-tap plot into quadrants.

With three quadrants defined, we can perform some basic statistical calculations on the data within each quadrant, such as means and standard deviations. For quadrants 1 and 3, we calculate the means and standard deviations of the magnitudes $(\bar{r}_1, \sigma_{r1}, \bar{r}_3, \sigma_{r3})$, rather than the x's and y's separately, because these quadrants contain data that are symmetric about the 45º axis. For quadrant 2, we calculate the means and standard deviations of the x's and y's separately, because this quadrant is on the off-diagonal. For the purpose of training our ANNs, we do not make use of the second quadrant's standard deviations, because they do not vary significantly with different combinations of impairments. One final parameter we make use of is similar to the Q-factor, which we define as $Q_{31} = (\bar{r}_3 - \bar{r}_1)/(\sigma_{r1} + \sigma_{r3})$.

To illustrate our method, we performed 125 simulations using the following impairment combinations: OSNR – 16, 20, 24, 28, and 32 dB; CD – 0, 200, 400, 600, and 800 ps/nm; and PMD with values of differential group delay (DGD) equal to 0, 10, 20, 30, and 40 ps. The simulated fiber channel included a laser with a center wavelength of

1550 nm and a FWHM line-width of 10 MHz; a 10 Gbit/s logic source; a single-arm, Mach-Zehnder optical modulator biased at the quadrature point with a V_π drive voltage; and a fourth-order Bessel-Thomson filter.

The ANN consisted of seven inputs ($\bar{r}_1, \sigma_{r1}, \bar{r}_3, \sigma_{r3}, \bar{x}_2, \bar{y}_2, Q_{31}$), three outputs (OSNR, CD, and DGD), and 28 hidden neurons. The ANN was trained by use of a software package developed by Zhang et al. [13]. Although alternatives were explored, a conjugate-gradient technique was chosen, because it offers a nice compromise in terms of memory requirements and implementation effort.

Once the model was trained, we validated its accuracy with a different set of testing data. We used 64 simulations with the following impairment combinations: OSNR – 18, 22, 26, and 30 dB; CD – 100, 300, 500, and 700 ps/nm; and DGD – 5, 15, 25, and 35 ps. The software reported a correlation coefficient of 0.97 for the testing data. Figure 4 compares the testing and ANN-modeled data for OSNR, CD, and DGD.

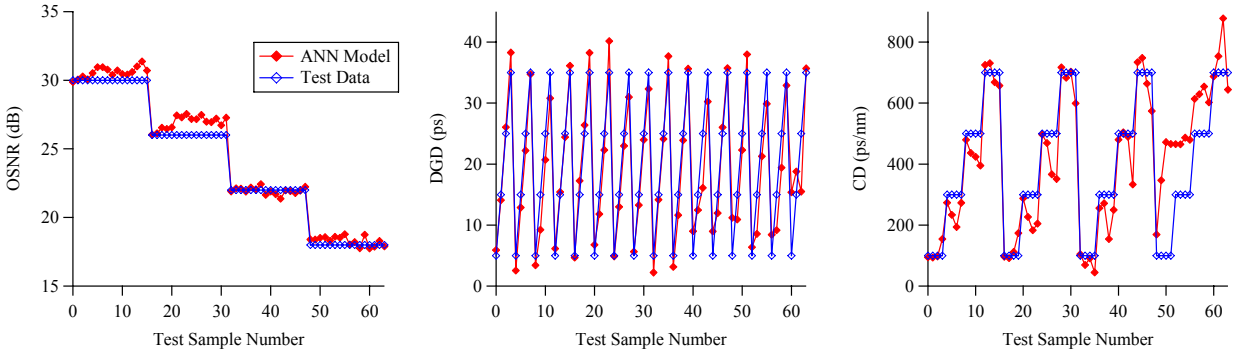

Fig. 4. Comparison of testing and ANN-modeled data for the 10 Gbps NRZ-OOK channel.

4. Conclusions

We have shown how ANN models trained with parameters derived from *B*/2 delay-tap plots can be used to simultaneously identify levels of OSNR, CD, and DGD for 10 Gbps NRZ-OOK signals. This method provides a powerful technique for monitoring the performance of optical channels without requiring synchronous sampling. Furthermore, the results of the ANN modeled in this paper compare favorably to those of the ANN developed by use of eye-diagram parameters in ref. 9.

5. References

[1] D. C. Kilper, R. Bach, D. J. Blumenthal, D. Einstein, T. Landolsi, L. Olstar, M. Preiss, and A. E. Willner, "Optical Performance Monitoring," *J. Lightwave Technol.*, vol. 22, no. 1, pp. 294-304, Jan. 2004.

[2] I. Shake, H. Takara, S. Kawanishi, and Y. Yamabayashi, "Optical Signal Quality Monitoring Method Based on Optical Sampling," *Electronics Letters*, vol. 34, no. 22, pp. 2152-2154, Oct. 1998.

[3] N. Hanik, A. Gladisch, C. Caspar, and B. Strebel, "Application of Amplitude Histograms to Monitor Performance of Optical Channels," *Electronics Letters*, vol. 35, no. 5, pp. 403-404, Mar. 1999.

[4] S. Ohteru and N. Takachio, "Optical Signal Quality Monitor Using Direct Q-Factor Measurement," *IEEE Photon. Technol. Lett.*, vol. 11, no. 10, pp. 1307-1309, Oct. 1999.

[5] S. D. Dods and T. B. Anderson, "Optical Performance Monitoring Technique Using Delay Tap Asynchronous Waveform Sampling," *OFC/NFOEC Technical Digest*, OThP5, Mar. 2006.

[6] S. D. Dods, T. B. Anderson, K. Clarke, M. Bakaul, and A. Kowalczyk, "Asynchronous Sampling for Optical Performance Monitoring," *OFC/NFOEC Technical Digest*, OMM5, Mar. 2007.

[7] B. Kozicki, A. Maruta, and K. Kitayama, "Asynchronous Optical Performance Monitoring of RZ-DQPSK Signals Using Delay-Tap Sampling," *ECOC 2007 Conference Proceedings*, P060, Sep. 2007.

[8] R. A. Skoog, T. C. Banwell, J. W. Gannett, S. F. Habiby, M. Pang, M. E. Rauch, and P. Toliver, "Automatic Identification of Impairments Using Support Vector Machine Pattern Classification on Eye Diagrams," *IEEE Photon. Technol. Lett.*, vol. 18, no. 22, pp. 2398-2400, Nov. 2006.

[9] J. A. Jargon, X. Wu, and A. E. Willner, "Optical Performance Monitoring Using Artificial Neural Networks Trained with Eye-Diagram Parameters," accepted for publication in *IEEE Photon. Technol. Lett.*

[10] M. H. Hassoun, "Fundamentals of Artificial Neural Networks," The MIT Press (1995).

[11] Q. J. Zhang and K. C. Gupta, "Neural Networks for RF and Microwave Design," Artech House (2000).

[12] K. Hornik, M. Stinchcombe, and H. White, "Multilayer Feedforward Networks Are Universal Approximators," *Neural Networks* (2), pp. 359-366, 1989.

[13] "NeuroModeler, ver. 1.5," Q. J. Zhang and His Neural Network Research Team, Department of Electronics, Carleton University, Ottawa, Canada, 2004.