Training of Neural Networks to Perform Optical Performance Monitoring of a Combination of Accumulated Signal Nonlinearity, CD, PMD, and OSNR

Xiaoxia Wu[1*], Jeffrey Jargon[2], Louis Christen[1], Alan Willner[1]
1. Dept. of Electrical Engineering - Systems, University of Southern California Email: xiaoxia@usc.edu
2. Department of Commerce, National Institute of Standards and Technology (NIST)

Abstract We propose a technique using artificial neural networks to simultaneously identify fiber nonlinearity, OSNR, CD, and PMD from eye-diagram and eye-histogram parameters. A correlation coefficient of 0.97 is obtained for a set of testing data.

Ⅰ. Introduction

High-performance optical networks are susceptible to various degrading effects that can change over time. Knowledge of the data channel degradation can be used to diagnose the network, repair the damage, drive a compensator/equalizer, and/or reroute traffic around a non-optimal link [1]. Therefore, it is valuable to monitor the channels for many types of impairments. Optical performance monitoring can be performed by measuring changes to the data and determining the "real-time" changes to various impairments, such that a change in a particular effect will change a measured parameter. This could employ: (i) optical techniques to monitor changes in an RF tone power or in the spectral channel power distribution [2], or (ii) electrical post-processing techniques in the specific case of using coherent detection [3].

Another type of optical performance monitoring is to probe the network upon initialization and train each receiver to record a specific data eye-diagram pattern that would correspond to a specified range of many physical parameters. Once the network is fully operational, variations of the received eye diagram from the ideal formation could then be attributed to specific physical parameters derived from the prior network/receiver training. Asynchronous scatter diagrams have been shown to give trends due to various impairments [4].

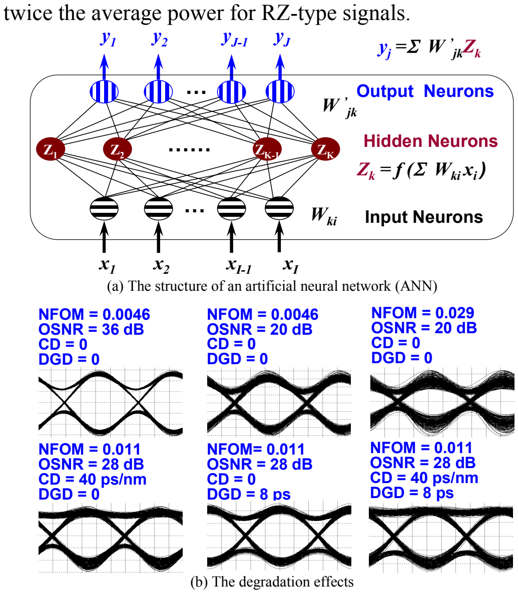
Recently, there have been reports of using a neural network approach [5, 6] and support vector machine (SVM) pattern recognition [7] to "train" the receivers in an optical network as to the relationship between resultant shapes of the data channel's eye diagrams and the degrading effects of optical signal-to-noise-ratio (OSNR), chromatic dispersion (CD), and first-order polarization-mode-dispersion (PMD) (i.e., differential group delay (DGD)). Importantly, the coefficients of the neural network algorithm are derived in a certain amount of iteration steps before live traffic is sent into the network.

One parameter that has not been explored in detail by use of the training approaches has been the accumulation of nonlinear impairments on the data channel. We emphasize that nonlinearity has typically been one of the most difficult parameters to monitor in an optical network.

In this paper, we get many sets of input-output data via network simulation in order to train a neural network as to the variations in the received eye diagrams due to accumulated channel nonlinear effects in addition to CD, DGD, and OSNR. Adding nonlinearity is a significant challenge in terms of the neural network approach, due to its specific signatures on the eye diagram. The number of inputs increases from 4 [5, 6] to 8. We show this technique in a 40 Gb/s return-to-zero (RZ) differential phase-shift-keying (DPSK) 3-channel WDM system, and a correlation coefficient of 0.97 is obtained for a set of testing data.

Ⅱ. Concept and Setup

Artificial neural networks (ANNs) are information-processing systems that learn from observations and generalize by abstraction [8]. 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, as shown in Fig.1 (a). ANNs learn the 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 reach prescribed values. In this case, the 4 outputs are nonlinearity figure of merit (NFOM), OSNR, CD, and PMD, and the 8 inputs include Q-factor, eye-closure, jitter, '0'-level crossing amplitude, mean of '1's, standard derivation of '1's, mean of '0's, and standard derivation of '0's. The NFOM is defined as $2\gamma PLBW_{3dB}$/channel spacing, with γ as the fiber nonlinearity coefficient, P as the average of the input optical powers of the neighboring channels that have the same spacing to the channel of interest, L as the fiber length, and BW_{3dB} as the signal bandwidth. After training, another set of data is used to test the ANN. The factor of '2' is used since the peak power is approximately twice the average power for RZ-type signals.



(a) The structure of an artificial neural network (ANN)

(b) The degradation effects

Fig. 1. Concept of ANN and the impact of degradation effects. NFOM: nonlinearity figure of merit, which equals to $2\gamma PLBW_{3dB}$ / channel spacing, with γ as the fiber nonlinearity coefficient, P as the average of the input optical powers of the neighboring channels, L as the fiber length and BW_{3dB} as the signal bandwidth.

The ANN architecture used here is a feed-forward, three-layer perceptron structure (MLP3) consisting of an input layer, a hidden layer, and an output layer. The hidden layer

978-1-4244-1932-6/08/$25.00 ©2008 IEEE

consists of 12 hidden neurons. The conjugate gradient method is used for training. Fig. 1 (b) shows some sampled eye diagrams with different impairments. We can clearly see that different impairment combinations imprint different signatures on the eye diagrams.

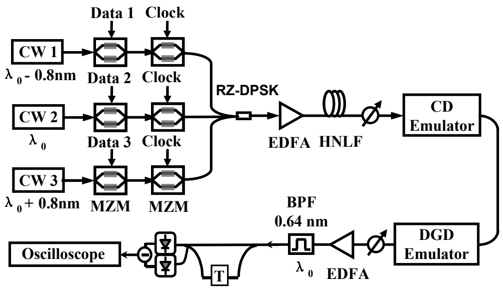


Fig. 2. Simulation setup. λ_0: the wavelength of the channel of interest. λ_0 = 1550 nm.

Fig. 2 shows the 3-channel WDM configuration used in the simulation. The 40 Gb/s RZ-DPSK signals are generated by two cascaded Mach-Zehnder modulators (MZM) and then coupled together with a channel spacing of 0.8 nm. The channels are decorrelated by use of different pseudo-random bit sequence (PRBS) orders. The WDM signals then pass through 2 km of highly nonlinear fiber (HNLF) with a nonlinear coefficient of 18 $W^{-1}km^{-1}$, zero dispersion wavelength of λ_0 (1550 nm), and dispersion slope of 0.05 $ps/nm^2/km$, following by a CD emulator and a PMD emulator. The output is sent to an Erbium-doped fiber amplifier (EDFA) with a variable optical attenuator in front to adjust the received OSNR. The signal is then filtered by a bandpass filter (BPF) with 0.64 nm bandwidth, and sent to an oscilloscope, where the eye diagram and eye histogram parameters are extracted. A 3-channel case is chosen to illustrate the concept, although this approach is also applicable to WDM networks with more channels.

Ⅲ. Results

The middle channel is chosen for the analysis because it experiences the strongest interchannel nonlinearity. The training data are obtained from the eye diagrams by use of a set of 135 samples (NFOM = 0.0046, 0.0072, 0.011, 0.018, 0.029; OSNR = 36, 28, 20 dB; CD = 0, 20, 40 ps/nm; DGD = 0, 4, 8 ps). The NFOM values correspond to -5 to 3 dBm input optical power in steps of 2 dB, respectively. Note that a few training samples are used in this work. In practical networks, a much larger amount of data will be required for training. Fig. 3 shows the training error versus epochs. An epoch is defined as a stage of ANN training that involves presentation of all the samples in the training data set to the neural network once for the purpose of learning [8]. The final training error is ~0.1 in our case.

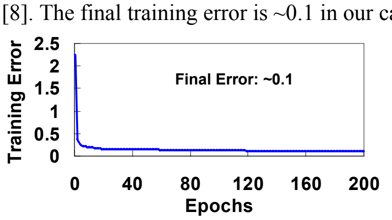
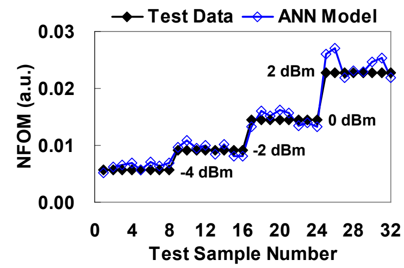
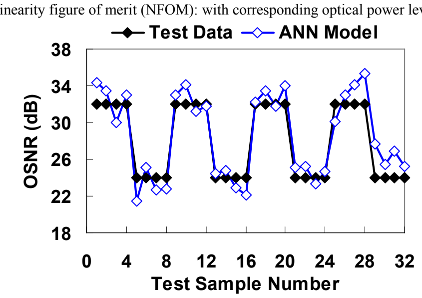


Fig. 3. Training error versus number of epochs for a 40-Gb/s RZ-DPSK channel (middle channel) in a 3-channel WDM system.

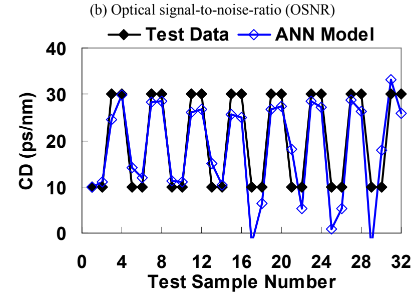
Once the model is trained, we validate its accuracy using a different set of testing data that includes 32 samples (NFOM = 0.0057, 0.0091, 0.014, 0.023; OSNR = 32, 24 dB; CD = 10, 30 ps/nm; DGD = 2, 4 ps). Again, the NFOM values correspond to -4 to 2 dBm input optical power in steps of 2 dB, respectively. The ANN reports a correlation coefficient of 0.97. Fig. 4 compares the testing and ANN-modeled data for NFOM, OSNR, CD, and DGD. It is shown that the ANN models, trained with parameters derived from eye diagrams and eye histograms, can potentially be used to simultaneously identify fiber nonlinearity, OSNR, CD, and PMD in WDM channels.



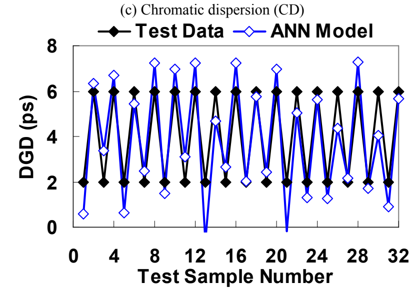
(a) Nonlinearity figure of merit (NFOM): with corresponding optical power level labeled



(b) Optical signal-to-noise-ratio (OSNR)



(c) Chromatic dispersion (CD)



(d) First-order polarization-mode-dispersion (PMD) (i.e., differential group delay (DGD))

Fig. 4. Comparison of testing and ANN-modeled data for a 40-Gb/s RZ-DPSK channel (middle channel) in a 3-channel WDM system.

Acknowledgements

This work was supported by the U.S. Department of Commerce and the DARPA-CORONET Program (BAA06-29), and is not subject to U.S. copyright. The authors thank Prof. Qi-Jun Zhang from Carleton University for providing the NeuroModeler software.

References

[1] D. C. Kilper, et al, JLT, 22(1), 2004.
[2] T. Luo, et al, OFC'03, ThY3, 2003.
[3] H. Sun, et al,Optics Express, 16(2), 2008.
[4] S. D. Dods, et al, OFC'06, OThP5, 2006.
[5] X. Wu, et al, ECOC'08, We.3.D.6, 2008.
[6] J. Jargon, et al, accepted to PTL, 2008.
[7] R. A. Skoog, et al, PTL, 18(22), 2006.
[8] Q. J. Zhang, et al, Neural Networks for RF and Microwave Design, Norwood, MA: Artech House, 2000.