Nonlinearity mitigation in polarization multiplexed fiber-optic transmission system based on fully-connected neural networks

S.A. Bogdanov[1,*], O.S. Sidelnikov[1], M.P. Fedoruk[1,2], S.K. Turitsyn[1,3]

[1]*Novosibirsk State University, Novosibirsk,*
[2]*Institute of Computational Technologies SB RAS, Novosibirsk,*
[3]*Aston Institute of Photonic Technologies, Aston University, Birmingham,*
[*]*E-mail: s.bogdanov@g.nsu.ru*

A growth of fiber-optic communication systems capacity by increasing the signal power (increasing signal-to-noise ratio in linear systems) is limited due to the nonlinear properties of optical fiber. Polarization-division multiplexing (PDM), which doubles the data transmission rate by using both polarizations, leads to additional nonlinear distortions caused by interactions between different polarizations. Machine learning approaches are powerful tools for digital signal processing in fiber-optic communications [1,2]. In this work we propose to use the fully-connected feed-forward neural networks (NNs) for nonlinearity mitigation at the receiver side of fiber-optic communication systems with polarization-division multiplexing. NNs of different architectures were studied and the comparison of their effectiveness was performed.

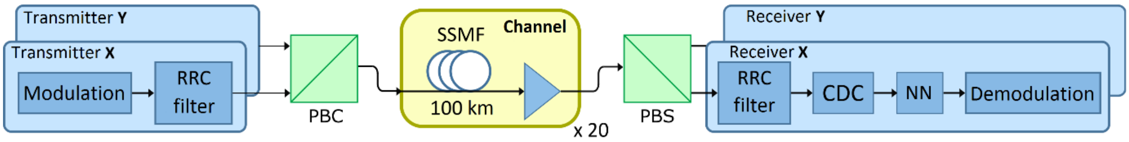

Fig. 1. Communication line scheme.

The scheme of the considered fiber-optic communication system is depicted in figure 1. Every transmitter generates 16-QAM signals with symbol rate R_s = 32 GBaud. The pulses are shaped using RRC filter with roll-off factor of 0.1. Then, signals from both transmitters are combined to one PDM-signal that is fed to the channel. The communication link consists of 20 spans of 100 km standard single mode fiber and erbium doped fiber amplifiers with NF = 4.5 dB after each span. At the receiver separated signal polarizations pass through a matched filter and, then, the ideal compensation of chromatic dispersion is performed. After this, the signal is downsampled to 1 sample per symbol and NN for nonlinear effects compensation is applied. Then, signal demodulation is performed and the bit error rate is calculated.

A neural network for 16-QAM symbols classification is used at the receiver. For every symbol processing the real and imaginary parts of N previous and N following symbols of both polarizations are fed to the input of NN. So, the input layer consists of $2\cdot2\cdot(2\cdot N+1)$ neurons. NN also has two hidden layers with a variable number of neurons and nonlinear activation function (tanh). Output layer consists of 16 neurons that is related to the number of points for the 16-QAM constellation. To train NN the Adam optimizer and TensorFlow are used.

In figure 2a results of two NN implementation at the communication system receiver are shown: first, the data is transmitted using one polarization state and, second, the data is transmitted using two polarizations but only one of them is fed to NN input (blue and green curve correspondingly). In the both cases for every symbol its 10 previous and 10 following neighbors were taken into account, and NN had 64 neurons on each hidden layer. Next, for two polarizations system the symbols from the both polarizations are used to predict symbols from the one of the polarizations (red dots). We can see that applying a NN with 64 neurons on hidden layers is ineffective because such simple networks are not able to process large numbers of symbols. But increasing the number of neurons in hidden layers lets to improve the effectiveness of using symbols from the second polarization.

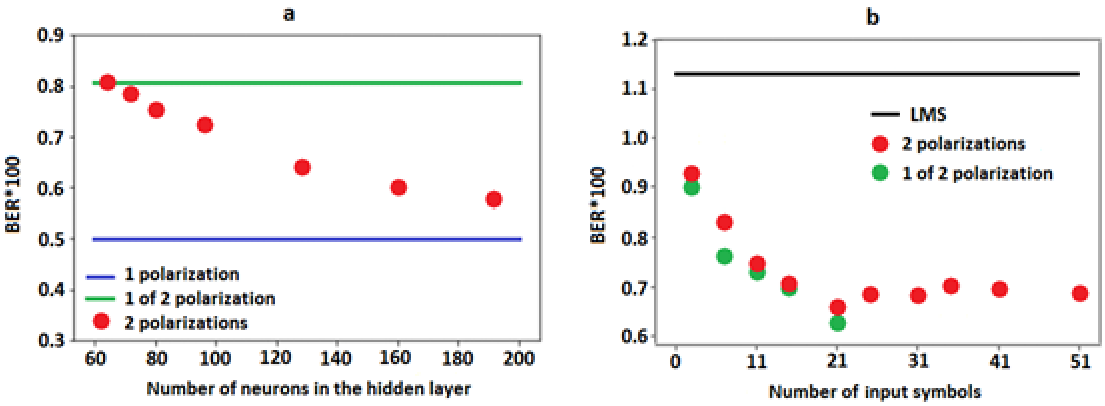

Fig. 2. Efficiency of accounting for the second polarization in dependence on number neuron in the hidden layer (a). The result of applying NN for prediction one and two polarizations (b).

Figure 2b depicts a result of data processing from both polarizations by the NN with 192 neurons in each hidden layer in dependence on the number of input symbols from each polarization. In the figure the result of a linear compensation scheme based on adaptive filters (Least Mean Square - LMS), that compensate for the nonlinear phase shift, is presented as a benchmark. In this case NN predicts data either from only one polarization or from both (green and red dots correspondingly). We see some degradation of data processing performance for two polarizations and this can be explained by the classification into 256 classes.

The work of O.S. Sidelnikov was supported by grant of the President of the Russian Federation (MK-915.2020.9). The work of S.A. Bogdanov was supported by the state funding program FSUS-2020-0034.

References

[1] O.S. Sidelnikov, et al, *Opt.Express.* **26**, 32765-32776 (2018).
[2] C. Häger, et al, *Proc. OFC*, W3A.4 (2018).