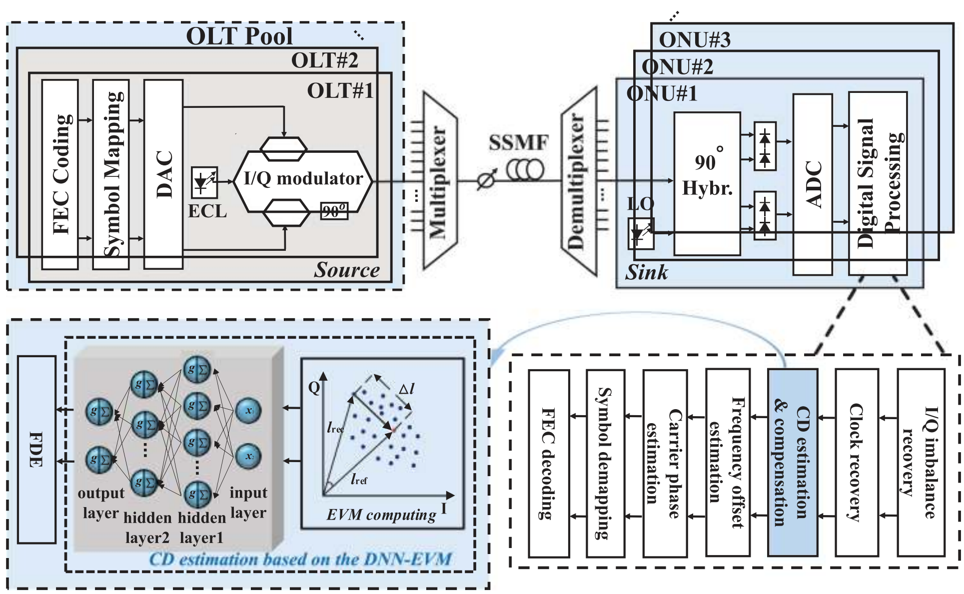
IEEE Photonics Journal
An IEEE Photonics Society Publication

Open Access

Low-Complexity Adaptive Chromatic Dispersion Estimation Scheme Using Machine Learning for Coherent Long-Reach Passive Optical Networks

Volume 11, Number 5, October 2019

Jin Li
Danshi Wang, *Member, IEEE*
Min Zhang



DOI: 10.1109/JPHOT.2019.2936426

IEEE photonics SOCIETY
IEEE

Low-Complexity Adaptive Chromatic Dispersion Estimation Scheme Using Machine Learning for Coherent Long-Reach Passive Optical Networks

Jin Li, Danshi Wang, *Member, IEEE*, and Min Zhang

State Key Laboratory of Information Photonics and Optical Communications, Beijing University of Posts and Telecommunications, Beijing 100876, China

DOI:10.1109/JPHOT.2019.2936426

Manuscript received July 17, 2019; revised August 14, 2019; accepted August 16, 2019. Date of publication August 21, 2019; date of current version August 30, 2019. This work was supported in part by the National Natural Science Foundation of China (NSFC) under Project 61705016, in part by the BUPT Excellent Ph.D. Students Foundation CX2019313, in part by the Fundamental Research Funds for the Central Universities under Grant 2019RC12, and in part by the National Key R&D Program of China under Grant 2016YFB0901200. Corresponding author: Danshi Wang (e-mail: danshi_wang@bupt.edu.cn).

Abstract: In the coherent long-reach passive optical networks (LRPON), it is crucial to propose cost-effective digital signal processing (DSP) technologies to reduce the overall complexity and power consumption. This paper has proposed a low-complexity chromatic dispersion (CD) estimation scheme based on deep neural networks (DNN) and the error vector magnitude (EVM). To add comparisons, the performances of CD estimation schemes using other two well-known machine learning algorithms including the k-nearest neighbor (KNN) and the decision tree (DT) have also been investigated. The simulation results show that the proposed CD estimation scheme is effective in the coherent LRPON with the quadrature phase shift keying (QPSK) and 16-ary quadrature amplitude modulation (QAM) systems at 14Gbaud rate, 28Gbaud rate and 56Gbaud rate. The comprehensive performances of the DNN outperform those of the KNN and the DT. The mean estimation error of the DNN is less than 20ps/nm within the 100 km access distance in the 28Gbaud QPSK/16QAM systems. What's more, compared with classical methods using the CD scanning and frequent domain equalizers (FDE), the computation complexity of the proposed CD estimation scheme based on the DNN-EVM has been respectively reduced by 72.3 times, 86.7 times and 2.8 times about the amount of multipliers, adders and comparators.

Index Terms: Chromatic dispersion estimation, coherent passive optical networks, machine learning, digital signal processing, low complexity.

1. Introduction

In the future fixed-mobile convergent passive optical network (PON), driven by emerging bandwidth-thirsty network services such as the high-definite video, the virtual/augmented reality and the cloud computing, the increasing capability demand between the optical line terminal (OLT) and the optical network unit (ONU) will be improved from 10Gbit/s to beyond 100Gbit/s [1]. Moreover, to increase the number of users and enlarge the area covering, the access distance is crucial to be extended, which is up to 100 kilometers [2]. To provide the beyond 100Gbit/s access capability and the long access reach, the coherent long-reach PON (LRPON) is considered as a promising candidate due to the high receiver sensitivity, frequency selectivity and capability scalability [3], [4]. For the coherent

PON, the digital signal processing (DSP) should be low-complexity to reduce the overall complexity and the power consumption of the ONU [5]. Recently, DSP technologies specially designed for the coherent PON have been rapidly developed to suit for the low-complexity requirements, where the simplified carrier recovery, the phase offset estimation and the adaptive equalization have been proposed [6]–[8].

The chromatic dispersion (CD) compensation is significant for the subsequent DSP modules in the coherent systems. The traditional CD compensation is based on the frequency domain equalizer (FDE) with fixed coefficients configured according to the given fiber parameters [9]. However, in the coherent LRPON, owing to the diverse access distances between the OLT and the multi-level ONU, the CD compensation through the FDE with the prior knowledge of fiber parameters including the distance and optical fiber coefficients is not adaptive enough. Thus, the accumulated CD is generally compensated by the adaptive equalization module where the access distance is typically within 20 km for the coherent PON [10]. However, in the coherent LRPON, the larger accumulated CD is beyond the compensation capability of the adaptive equalizer. Therefore, adaptive CD compensation methods with larger dynamic ranges should be proposed for the coherent LRPON.

In the metro and backbone networks, the adaptive CD compensation methods have the large dynamic range and they are generally based on the CD scanning and the FDE, where varied CD values with certain scanning steps are sent into the FDE module until the peak-to-average-power ratio (PAPR), the value of the Godard error function or the parameter extracted from the delay-tap plot of the received signal reach to the minimum value [11]–[13]. These adaptive CD compensation methods are straightforward and effective. However, they are more suitable for the CD compensation in the metro and backbone network, where the sufficient DSP resources are available in coherent receiver to support the frequent scanning and complex fast Fourier transformation (FFT) operations. Complied with the low-complexity demands of the ONU, more hardware-efficient adaptive CD compensation methods should be proposed for the DSP in the coherent LRPON.

In this paper, to reduce the computation complexity, we propose a low-complexity adaptive CD compensation scheme for the coherent LRPON, where the frequent linear scanning and the FDE process are substituted by the proposed CD estimation method. Recently, machine learning (ML) has been successfully applied in the optical networks [14]–[19], which is specialized in modeling complicated functional relationships where the underlying principles of physic and mathematics are difficult to be described [20], [21]. It is observed that the error vector magnitude (EVM) of the received signal after the in-phase and quadrature (I/Q) balance and the clock recovery monotonically increases with the accumulated CD within the access span in the coherent LRPON, which enlightens us to adopt ML algorithms to discover the hidden functional relationship between the EVM and the CD. Further, the measured EVM of the received signal is sent to the ML module and then the estimated CD is fed into the equalizer for the CD compensation. Further, the performances of the EVM-ML based CD estimation method are evaluated in the typical quadrature phase shift keying (QPSK) and 16-ary quadrature amplitude modulation (QAM) systems with 14Gbaud rate, 28Gbaud rate and 56Gbaud rate. Three typical ML algorithms including the deep neural network (DNN), the k-nearest neighbor (KNN) and the decision tree (DT) are adopted. The simulation results show that the comprehensive performances of the DNN outperform those of the KNN and the DT and the average CD estimation error of the DNN is less than 20ps/nm within the 100 km access distance in the 28GBaud QPSK/16QAM coherent systems between the OLT and the ONU. To add comparisons, other classical DNN-based CD estimation methods, i.e. the DNNs trained with the eye-diagram parameters (EDP) [22], [23], the asynchronous amplitude histograms (AAH) [24], the asynchronously sampled signal amplitudes (ASSA) [25], the balanced-detected asynchronous diagrams (BDAD) [26] and the asynchronous constellation diagrams (ACD) [27], have also been compared with the proposed DNN-based approach about the estimation accuracy, the dynamic range, the network structure and the applicable systems. The results show that larger CD estimation range, more advanced coherent systems and the similar estimation accuracy are available in the proposed DNN-based approach and it is more suitable for the low-complexity CD estimation for the coherent LRPON. What's more, compared with the classical CD scanning and FDE methods, the computation complexity of the proposed CD estimation method has been decreased dramatically. The computation complexity of the proposed CD estimation scheme based on the DNN has been respectively reduced by 72.3 times, 86.7 times and 2.8 times in terms of the amount of multipliers, adders and comparators compared with the FDE-scanning method in the typical case. The proposed CD estimation scheme based on the EVM-DNN has the potential to be applied in the hardware-efficient DSP of the ONU in the coherent LRPON.

2. Principal of the Proposed Scheme

The EVM is an important index indicating the comprehensive impairments of the received signal [28]. In the coherent optical communications, the value of EVM of the received signal after the I/Q balance and the clock recovery is sensitive to the accumulated CD within certain ranges. Generally, when the transmitted optical signal experiences less CD, the deviation degree of the received signal constellation points from the ideal constellation points is small accordingly. In contrast, the signal constellation is greatly dispersed while the CD value becomes larger. Therefore, the EVM of the received signal obtains the CD information and it is strongly correlated with the accumulated CD. In the coherent LRPON, through analyzing sufficient samples of the EVM of the received signal experiencing various accumulated CD after the balance detection, I/Q imbalance recovery and time clock recovery, the hidden relationship between the EVM and the corresponding accumulated CD can be researched, which has the potential to be used to estimate the value of the CD for the coherent LRPON. The value of EVM is defined as:

$$EVM = \sqrt{\frac{\frac{1}{S}\sum_{s=1}^{S}((I_s - I'_s)^2 + (Q_s - Q'_s)^2)}{\frac{1}{S}\sum_{s=1}^{S}(I'^2_s + Q'^2_s)}} \tag{1}$$

Where S denotes the size of the sample dataset. (I_s, Q_s) and (I'_s, Q'_s) refer to the $s-th$ constellation point of the practical received signal and its corresponding constellation point of the ideal signal respectively. To estimate the value of the accumulated CD accurately and efficiently, the specific mapping relationship between the value of EVM and the accumulated CD is constructed through the classic machine learning algorithm called DNN. The DNN is selected due to its powerful nonlinear mapping ability where the simple one-hidden-layer DNN is capable of fitting any nonlinear continuous function with arbitrary accuracy [29], [30].

The procedure of the proposed DNN is made up of two basic procedures including the forward propagation and the back propagation. The original input vector $x = (x_1, x_2, ..., x_N)$ is firstly combined with weighted input $z_i^{(l+1)}$ linearly in the neurons, where the superscript $(l+1)$ indicates that the corresponding parameters are in the $(l+1)-th$ layer of the DNN. The N_l denotes the number of nodes in the $l-th$ layer and $a_j^{(l)}$ refers to the original input of the $j-th$ neuron. Moreover, the parameters $W_{ij}^{(l)}$ and $b_i^{(l)}$ indicate the weights and biases in the $l-th$ layer respectively. Further, the weighted input is fed into the nonlinear activation function $g(z)$ and the output a_j is known as the activation.

$$a_i^{(l+1)} = g\left(\sum_{j=1}^{N_l}\left(W_{ij}^{(l)} \cdot a_j^{(l)} + b_i^{(l)}\right)\right) \tag{2}$$

In the forward propagation, the input is transformed nonlinearly by the activation functions. Typical activation functions consist of the hyperbolic tangent (tanh) function, softplus function and sigmoid function. These functions need to be nonlinear and differential to increase the nonlinearity of the transformation and enable the optimization of the DNN. The selection of the activation function has significant effects on the nonlinear mapping capacity and convergence ability of the specific DNN. Moreover, taking the single-hidden-layer DNN for example, the output of the DNN $\overline{y}$ is described in (3) and then the difference between the predicted output $\overline{y}$ and the real output y is calculated

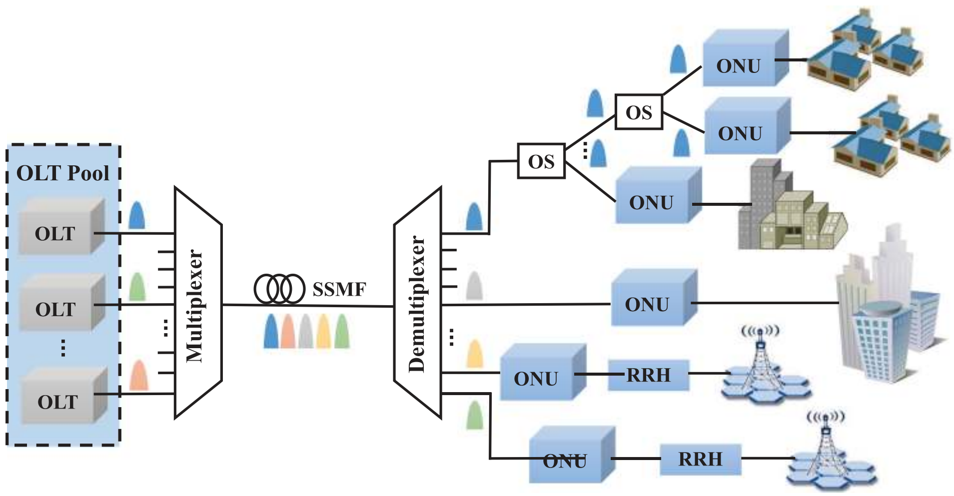


Fig. 1. The architecture of the fixed-mobile convergent coherent LRPON. LRPON: long-reach passive optical networks; OLT: optical line terminal; ONU: optical network unit; RRH: remote radio head; OS: optical splitter; SSMF: standard single mode fiber.

through the cross entropy loss function defined in (4):

$$\overline{y}_{W,b}(x) = g\left(\sum_{i=1}^{N_2} W_{ki}^{(2)} \cdot g\left(\sum_{j=1}^{N_1} W_{ij}^{(1)} \cdot x_j + b_i^{(1)}\right) + b_k^2\right) \tag{3}$$

$$L\left(W, b\right) = -\frac{1}{m}\left(\sum_{i=1}^{m} y^{(i)} \log \overline{y(x^{(i)})} + (1 - y^{(i)}) \log(1 - \overline{y(x^{(i)})})\right) + \frac{\varsigma}{2}\sum_{l=1}^{N_3}\sum_{i=1}^{N_2}\sum_{j=1}^{N_1}\left(W_{ij}^{(l)}\right)^2 \tag{4}$$

Where m means the size of the training dataset. The parameter ς used in the regularization term denotes the weight of connectivity weight parameter $W_{ij}^{(l)}$, which tends to enable $W_{ij}^{(l)}$ to be zero to simplify the structure of the network and then increase the generalization capability of the DNN. After calculating the loss function, the objective of the back propagation is to update the parameters $W_{ij}^{(l)}$ and $b_i^{(l)}$ based on the batch gradient descent (BGD) method, where the gradient of the loss function in terms of $W_{ij}^{(l)}$ and $b_i^{(l)}$ is fully utilized to improve the speed with which the minima of the loss function can be located. The principle of the BGD is shown in (5):

$$\begin{cases} W_{ij}^{(l)} = W_{ij}^{(l)} - \beta \frac{\partial}{\partial W_{ij}^{(l)}} L\left(W, b\right) \\ b_i^{(l)} = b_i^{(l)} - \beta \frac{\partial}{\partial b_i^{(l)}} L\left(W, b\right) \end{cases} \tag{5}$$

Where β is the learning rate. During the update, total training dataset is included in the loss function and the gradients of $W_{ij}^{(l)}$ and $b_i^{(l)}$. During each iteration, the weight matrix $W_{ij}^{(l)}$ and the bias matric $b_i^{(l)}$ are moved towards the direction of the greatest rate of deceasing the value of the loss function. The parameter matrix $W_{ij}^{(l)}$ and $b_i^{(l)}$ will be determined until the value of the loss function is not more than O^{eror}, the error margin approximating zero and it is set as 0.0001 in this paper. Finally, the well-trained DNN is capable of constructing the nonlinear mapping relationship between the input and output variables. In the proposed adaptive CD estimation scheme based on the DNN, diverse samples consisting of varied CD and the corresponding EVM of the received signal are gathered as the dataset and then fed into the DNN. After the training process, the trained DNN can establish the mapping relationship between EVM and the CD. Once the EVM of the received signal is measured, the corresponding value of the accumulated CD is estimated by the well-trained DNN as shown in

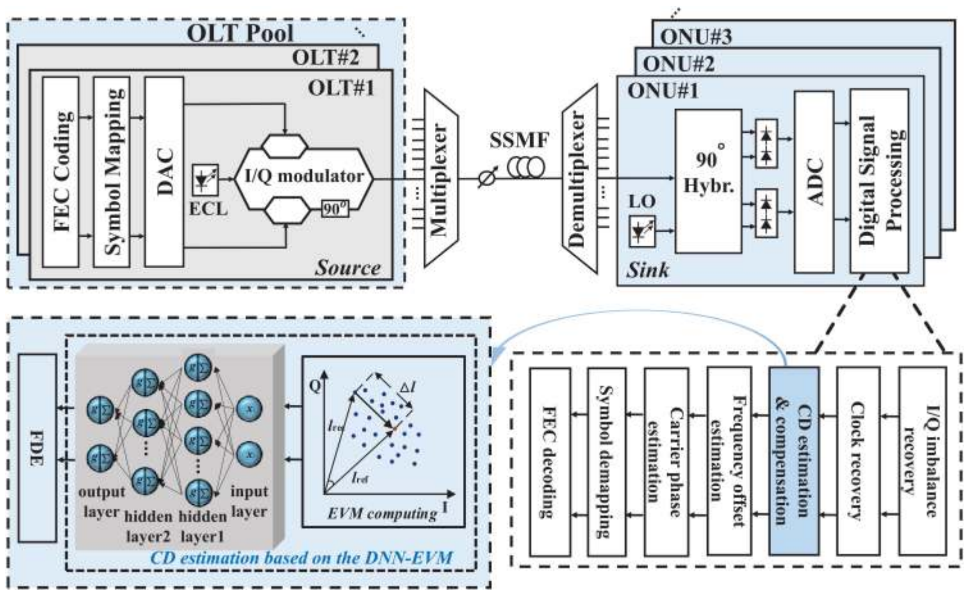


Fig. 2. The schematic diagram of the proposed chromatic dispersion estimation scheme based on the DNN-EVM for the coherent LRPON. DNN: deep neural network; EVM: error vector magnitude; LRPON: long-reach passive optical networks.

the Fig. 2. Finally, the estimation error can be measured through comparing the estimated CD with the real CD of the corresponding received signal in the coherent LRPON.

3. Simulation Results and Discussion

To evaluate the performances of the proposed CD estimation scheme, 14Gbaud, 28Gbaud and 56Gbaud QPSK/16QAM simulation platforms are constructed through the VPI transmission 8.6. In the Fig. 1, taking the 16QAM coherent systems between the source in the OLT and the sink in the ONU for example, the transmitter and the local oscillator laser work at 1550 nm with the typical 100 kHz linewidth. What's more, the input power of the transmitter is set as 0 dBm. Firstly, the pseudo-random binary sequence (PRBS) with the length of $2^{16}-1$ is mapped into 16QAM symbols, which are further sent into the digital-to-analog converter (DAC) module. After the fourfold oversampling, the in-phase and quadrature information of the 16QAM symbols is transformed into the corresponding electrical signal, which modulates the I/Q modulator consisting of two Mach-Zehnder modulators (MZM) with an additional 90° phase shift in one branch. Further, the modulated optical signal is coupled into the standard signal mode fiber (SSFM) with the varied length ranging from 5 km to 125 Km and the specific parameters of the SSMF are listed below: loss coefficient $\alpha = 0.2\,\mathrm{dB/km}$, dispersion coefficient $D = 16\,\mathrm{ps/(nm \cdot km)}$ and the nonlinear coefficient $\gamma = 1.3\,\mathrm{W^{-1}Km^{-1}}$.

At the coherent receiver, the received optical signal is mixed with the local oscillator (LO) in the 90° optical hybrid and the mixed signal is detected by the balanced photo-detectors. In the offline DSP, the resampling, I/Q imbalance recovery and time clock recovery are carried out and then the EVM of the received signal is calculated and collected for the DNN-EVM-based CD estimation module. For every value of CD at the step of 40 ps/nm ranging from 0ps/nm to 2000ps/nm, 30 sets of the EVM of the received signals are gathered and labelled with the corresponding CD class. In other words, there are 50 categories of EVM values from the received signals experiencing the diverse accumulative CD ranging from 0 to 2000 ps/nm and the corresponding CD labels are varied from 1 to 50. After the simulation, 1500 sets of data are collected, where 2/3 of them serve as training sets and others are divided into the testing sets. Further, these data sets are fed into the adaptive CD estimation module to train the DNN and then test the performances of the

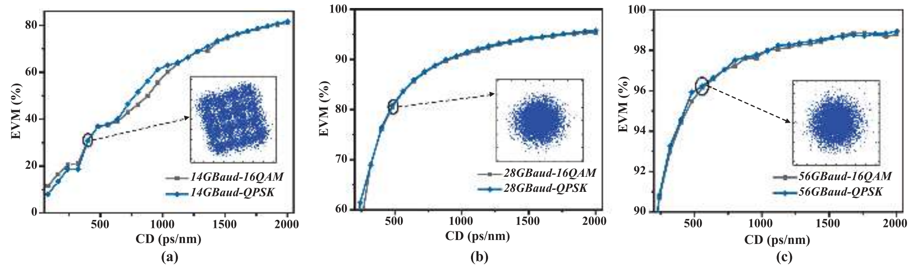


Fig. 3. The EVM curves in terms of the varied accumulated chromatic dispersion under the cases of (a) 14Gbaud, (b) 28Gbaud and (c) 56Gbaud coherent optical communication systems between the OLT and the ONU in the coherent LRPON.

trained DNN. During the training phase, the DNN is trained to discover the mapping relationship between the CD labels and the EVM, where appropriate parameters and structures of the DNN are determined after some iterations. During the testing phase, the EVM of the received signal after the I/Q balance and the clock recovery is calculated and then the CD class is estimated by the trained DNN. The estimation error can be calculated by comparing the true CD with the estimated CD to evaluate the performances of the DNN-EVM-based adaptive CD estimation scheme. Moreover, the CD estimation accuracy performance, the key parameters and the structure of the DNN are investigated as follows.

3.1 The CD Estimation Accuracy Performance

After the I/Q balance and the clock recovery, we measure the EVM of received signal experiencing the varied CD ranging from 0 ps/nm to 2000 ps/nm. The EVM curves in terms of different values of the CD are depicted in the Fig. 3. As shown in the Fig. 3, the value of EVM monotonically increases with the accumulated CD. In the 28Gbaud-QPSK optical coherent communication system, with the increase of the accumulated CD, the signal constellation points become more and more disperse and the corresponding EVM of the received signal is augmented from 19.2% to 97.6%. Moreover, the EVM curves in the coherent optical communication systems working at the same baud rate are almost overlapped, which means that the mapping relationship between the EVM and the CD is relatively transparent to the widely used modulation formats including the QPSK and the 16QAM. However, the hidden mathematical relationship between the EVM and the CD are complicated to be numerically analyzed. Therefore, the self-learning DNN algorithm is adopted to learn the functional mapping relationship between them and provide a reasonable prediction of the CD value for the collected EVM of the received signal derived from the coherent LRPON.

To research the specific mapping between the EVM of the received signal and the accumulated CD, the classic DNN is adopted. The reasons why the DNN is selected are that the DNN is capable of establishing any nonlinear function relationship with arbitrary accuracy. In the DNN, different activation functions, number of neurons in the hidden layers, number of hidden layers and the learning rate have significant effects on the CD estimation performances and the comparison results are shown in the Fig. 4. With the increase of the number of iterations, the CD estimation performance of the DNN is convergent gradually. Taking the 28GBaud-16QAM systems for example, after 500 iterations, the mean CD estimation error is gradually convergent and the mean error of the DNN is near 11 ps/nm. In the Fig. 4(a), the effect of different activation functions is investigated and the mean CD estimation error of the DNN with the sigmoid function has faster convergent speed and more stable estimation accuracy. Therefore, the sigmoid function is selected as the activation function. Further, the number of neurons in the hidden layers and the number of hidden layers of the DNN need to be specified. Seen from the Fig. 4(b), on the one hand, with the increase of the amount of hidden neurons and hidden layers, the learning capability of the DNN is improved and the convergent mean estimation error is decreased from 97.3 ps/nm to 11.0 ps/nm. On the other hand, due to the overfitting, the estimation error increases when the structure of the DNN becomes

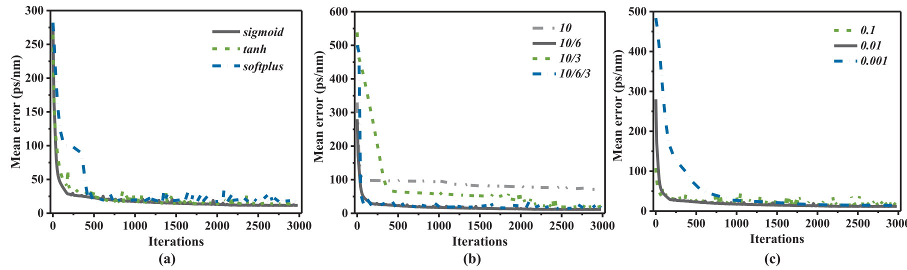


Fig. 4. The performances of the proposed CD estimation scheme based on the DNN with different (a) activation functions; (b) number of hidden neurons and hidden layers (10/6/3 means that there are respectively 10, 6 and 3 neurons in the first hidden layer, the second hidden layer and the third hidden layer); (c) learning rates (ranging from 0.001 to 0.1) in the 28Gbaud-16QAM coherent systems.

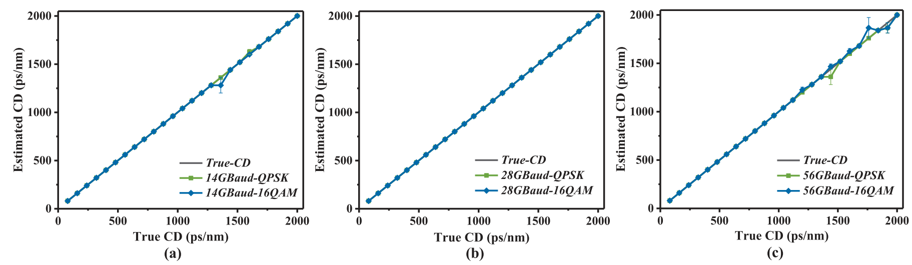


Fig. 5. The CD estimation performances of the proposed CD estimation scheme based on the DNN for the coherent optical communication systems with (a) 14Gbaud, (b) 28Gbaud and (c) 56Gbaud rate.

more complex. Thus, appropriate number of hidden neurons and hidden layers are important and two hidden layers are adopted due to the good balance between the learning capability and the generalization performance, where the number of neurons is set as 10 and 6 in the first hidden layer and the second hidden layer respectively. Moreover, the influence of various learning rates is also researched. As displayed in the Fig. 4(c), the faster convergent speed is available but larger fluctuation is also existed when the learning rate is too large. To keep the balance of the convergent speed and the stable estimation accuracy, 0.01 is chosen as the learning rate in the proposed DNN.

Further, the CD estimation accuracy performance curves of the DNN are investigated for the QPSK and 16QAM coherent optical communication systems working at 14GBaud, 28GBaud and 56GBaud rate respectively. It can be seen from the Fig. 5 that the estimated CD is generally corresponded with the true CD, which means that the estimation error is much small in the coherent systems with 14Gbaud, 28Gbaud and 56Gbaud rate. To add comparisons, the adaptive CD estimation schemes based on other two classical machine learning algorithms, i.e. KNN and DT, are also evaluated. In the comparisons, the key parameters are selected through the grid search and cross validation. In the KNN, the significant parameter needs to be chosen is the value of k, the number of nearest neighbors, is set as 8 and 12 respectively for the QPSK and 16QAM systems. Moreover, in the DT, the important parameter is the maximum number of splits and this parameter is valued in 150. In the Fig. 6(b), the mean CD estimation accuracy performance of the adaptive CD estimation method based on the DNN is superior to those of the KNN and the DT in the case of lager CD range generally.

In the 28GBaud-QPSK systems, the mean CD estimation error is around 4.5 ps/nm, which is less than the error of the KNN and the DT slightly. However, with the increase of the baud rate and the modulation efficiency, the estimation accuracy of the DNN-EVM-based method obviously exceeds those of the KNN and the DT in general. For the 56GBaud-16QAM systems, the mean error of the CD estimation method based on the DNN is near 51.9 ps/nm, which is less 16.4% and 23.7% than those of the KNN and the DT respectively. The reasons why the CD estimation method based on

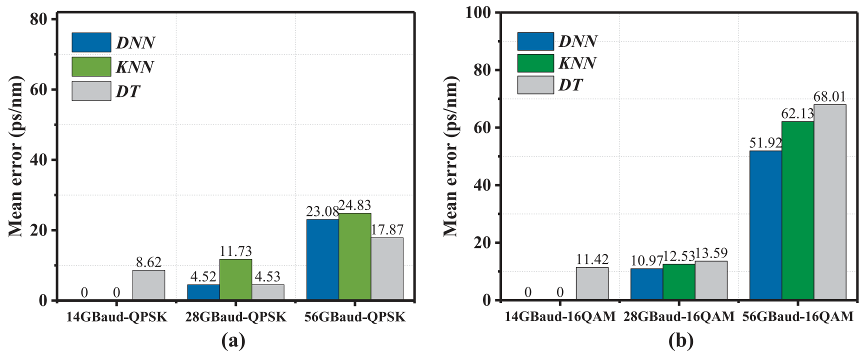


Fig. 6. The estimation error of the CD estimation schemes based on the DNN, the KNN and the DT algorithms for (a) the QPSK systems and (b) the 16QAM systems when the CD ranges from 0ps/nm to 2000ps/nm in the coherent LRPON.

TABLE 1
The Complexity Comparisons Among Different CD Estimation Schemes

Scheme type	Multipliers	Adders	Comparators
FDE-scanning	$2MN \cdot \log_2 N + 5MN + 4M$ (451150)	$3MN \cdot \log_2 N + 2MN$ (716800)	$M(M-1)/2$ (45)
DT-EVM	$3N + 3$ (6147)	$4N$ (8192)	$C - 1$ (49)
KNN-EVM	$3N + N_T + 3$ (7147)	$4N + N_T$ (9192)	$N_T \cdot \log_2 N_T$ (9967)
DNN-EVM	$3N + N_1N_2 + N_2N_3 + N_3N_4$ (6236)	$4N + (N_1-1)N_2 + (N_2-1)N_3 + (N_3-1)N_4$ (8266)	$N_2 + N_3$ (16)

the DNN outperforms those of the KNN and the DT are that the DNN is capable of constructing essential nonlinear mapping with high accuracy and building the appropriate model to describe the correlative relationship between variables. The well trained DNN is capable of providing the reasonable prediction for the EVM of the unforeseeable received signal and becomes less sensitive to the random noise. Comparatively, the KNN and the DT are intelligent and simple but they are not good at discovering the intrinsic relationship between the input and the output.

3.2 The Computation Complexity Analysis

Besides the CD estimation accuracy, the computation complexity performances of the CD estimation schemes are critical to be investigated for the practical applications especially for the coherent LRPON. Next, the specific computation efficiency performances of the traditional FDE-scanning method and adaptive CD estimation methods based on the DNN, the KNN and the DT will be researched. In the classical CD estimation methods based on the FDE and the CD scanning for the metro and core networks, different value of the CD with the certain scanning step is sent into the FDE until the minimum value of the features of the received signal is obtained. The frequent scanning and numerous FFT and IFFT operations are required. For one CD scanning, the received signal is processed through the FFT, the FDE and the IFFT. In each FFT/IFFT operation, the number of real multiplications and real additions is $2N \cdot \log_2 N$ and $3N \cdot \log_2 N$ respectively, where N denotes the number of samples in the block and is also taken as the FFT size. In the FDE, $4N + N + 4$ real multiplications and $2N$ real additions are needed for the complex multiplication in the frequency domain. Given the number of the CD scanning is M, the total multipliers, adders and the comparators are listed in the Table 1.

Moreover, for the proposed adaptive CD estimation schemes based on the machine learning and the EVM, the computation complexity mainly consists of two parts. In the one part, $3N + 3$ real

TABLE 2
The Processing Time Comparisons Among Different ML Algorithms

Algorithm	DNN	KNN	DT
Training time	$0.233ms/epoch$	$0ms$	$0.331ms$
Testing time	$0.129ms$	$32.217ms$	$0.162ms$

multiplications and $4N$ real additions are required for the EVM computation according to (1). In the other part, the computation complexity is constituted with the training complexity and the testing complexity in the machine learning, where the training procedure can be accomplished offline and the testing complexity is regarded as the major computation complexity of the machine learning enabled applications. In the DNN, the number of neurons in the input layer, the first hidden layer, the second hidden layer and the output layer is respectively N_1, N_2, N_3 and N_4. When the DNN is well trained, $N_1N_2 + N_2N_3 + N_3N_4$ multipliers, $(N_1 - 1)N_2 + (N_2 - 1)N_3 + (N_3 - 1)N_4$ adders and $N_2 + N_3$ comparators are needed. In the KNN, Euclidean distances among the testing sample and all training samples are calculated and then the first k, the number of nearest neighbors, distances in the ascending order are obtained, where the number of multipliers, adders and comparators is N_T (the size of the training samples), N_T and $N_T \cdot \log_2 N_T$ respectively. In the DT, $C - 1$ comparators are required during the decision process, where C denotes the number of CD value categories. For a fixed $N = 2048$ and $M = 10$ (typical scanning times), the required number of multipliers, adders and comparators for the different CD estimation schemes are concluded in the Table 1. The results show that the complexity of the classical CD estimation methods is greatly increased with the block length N and the number of CD scanning M. Compared with the FDE-scanning method, the computation complexity of the proposed CD estimation scheme is decreased dramatically. As shown in the Table 1, for the DNN-EVM-based CD estimation method, the amount of multipliers, adders and comparators is respectively reduced by 72.3 times, 86.7 times and 2.8 times compared with the FDE-scanning method in the typical case. Therefore, the proposed low-complexity CD estimation method is more suitable for the DSP of the ONU in the coherent LRPON.

Further, the processing time of different ML algorithms is also investigated. The processing time of the proposed ML-based CD estimation approaches consists of the training time and the testing time. The specific processing time relies on the processors. In this work, the typical central processing unit (CPU), i.e. Intel(R) Core(TM) i7-6700 CPU @3.40 GHz, is utilized to calculate the processing time of different ML-based CD estimation methods. Since the machine learning algorithm can be trained off line generally, the time consumption for the proposed ML-based CD estimation is mainly the testing time. As displayed in the Table 2, the testing time of the DNN is less than those of the KNN and the DT. The specific testing time is 0.129 ms, 32.217 ms and 0.162 ms in the DNN, the KNN and the DT respectively. In the DNN, the testing time is similar to that of the DT and far less than that of the KNN, which is respectively 79.6% and 0.4% of those of the DT and the KNN. This is consistent with the complexity comparisons in the Table 1 and the testing time of different ML algorithms increases with the computation complexity.

In the coherent LRPON, the high CD estimation accuracy and low complexity are both significant to guarantee the CD compensation performances and reduce the overall complexity of the ONU respectively. In the Table 1, the computation complexity of the DNN is similar to that of the DT, which is far less that of the KNN. Moreover, the mean CD estimation accuracy of the DNN exceeds those of the KNN and the DT generally. In the 14GBaud-QPSK systems, the zero CD estimation error is available in the DNN while the mean estimation error is 8.62ps/nm in the DT. For the 56GBaud-16QAM systems, the mean error of the CD estimation method based on the DNN is near 51.9ps/nm, which is 23.7% less than that of the DT. Therefore, after the comprehensive consideration about the estimation accuracy and the computation complexity, the DNN is chosen for the CD estimation in the coherent LRPON. To improve the completeness of discussions, other classical DNN-based CD estimation methods including the DNNs trained with the EDP, the AAH, the ASSA, the BDAD and the ACD will be compared with the proposed DNN-based approach in terms of the mean estimation

TABLE 3
The Comparisons Among Different DNN-Based CD Estimation Approaches

Approach	Mean error	Dynamic range	Network structure	Systems
Proposed	$0-51.92ps/nm$	$0-2000ps/nm$	$1/10/6/1$	$28/56/112Gbps\ QPSK$ $56/112/224Gbps\ 16QAM$
EDP-DNN [22]	$3.98-4.56ps/nm$	$0-60ps/nm$	$4/12/3$	$40Gbps\ OOK$ $40Gbps\ DPSK$
EDP-DNN [23]	$9.10-47.44ps/nm$ $(RMSE)$	$0-800ps/nm$	$24/40/3$	$10/20Gbps\ OOK$ $40Gbps\ QPSK$
AAH-DNN [24]	$9.66-9.82ps/nm$ $(RMSE)$	$0-400ps/nm$	$200/326/3$	$40Gbps\ DQPSK$ $40Gbps\ 16QAM$
ASSA-DNN [25]	$17.00-29.00ps/nm$ $(RMSE)$	$-500-500ps/nm$	$10/42/3$	$40/56Gbps\ QPSK$ $40Gbps\ DPSK$
BDAD-DNN [26]	$8.75ps/nm$ $(RMSE)$	$0-50ps/nm$	$7/12/3$	$100Gbps\ QPSK$
ACD-DNN [27]	$18.71ps/nm$ $(RMSE)$	$0-200ps/nm$	$7/28/3$	$40Gbps\ QPSK$

error, the dynamic estimation range, the network structure (the number of neurons in the input layer, the hidden layer/layers and the output layer) and the applicable systems. It can be seen from the Table 3 that the dynamic CD estimation range of the proposed method (0–2000 ps/nm) obviously exceeds those of other classical DNN-based approaches and typical high-speed coherent systems (beyond 100Gbps QPSK/16QAM) are also covered. Moreover, the CD estimation error of the proposed DNN-based method is similar to those of the classical DNN-based approaches with the larger dynamic range. The mean estimation error increases with the CD estimation range generally. In the proposed method, the CD estimation range is 0–2000 ps/nm while the mean estimation error is between 0ps/nm and 51.92 ps/nm. In the [23], the root mean square error (RMSE) ranges from 9.10 ps/nm to 47.44 ps/nm, which is close to that of the proposed DNN-based method. However, the dynamic range is within 800 ps/nm, which is 40% of that of the proposed method. Due to larger CD estimation ranges, more advanced coherent systems and similar estimation accuracy, the proposed DNN-based approach is more suitable for the low-complexity CD estimation for the coherent LRPON.

4. Conclusions

We have proposed a low-complexity CD estimation scheme based on the DNN-EVM to decrease the computation complexity of the ONU in the coherent LRPON. To add comparisons, the performances of CD estimation schemes using the KNN and the DT have also been evaluated. The simulation results show that the proposed CD estimation scheme is effective in the coherent LRPON with the QPSK and 16QAM coherent systems working at 14Gbaud rate, 28Gbaud rate and 56Gbaud rate. The estimation accuracy performances of the DNN exceed those of the KNN and the DT with relatively small-size computation complexity. The mean estimation error of the DNN is less than 20ps/nm within the 100 km access distance in the 28Gbaud QPSK/16QAM transmission systems between the OLT and the ONU in the coherent LRPON. Further, six other classical DNN-based CD estimation methods have also been compared with the proposed DNN-based approach. The results show that larger CD estimation range, more advanced coherent systems and the similar estimation accuracy are available in the proposed DNN-based approach. Moreover, compared with the classical CD scanning and FDE methods, the computation complexity of the CD estimation scheme based on the DNN-EVM has been reduced by 72.3 times, 86.7 times and 2.8 times about the number of multipliers, adders and comparators respectively in the typical case. The proposed CD estimation scheme has the potential to be one of the members of hardware-efficient DSP technologies for the coherent LRPON.

References

[1] A. Shahpari *et al.*, "Coherent access: A review," *J. Lightw. Technol.*, vol. 35, no. 4, pp. 1050–1058, Feb. 2017.

[2] D. Lavery, R. Maher, D. S. Millar, B. C. Thomsen, P. Bayvel, and S. J. Savory, "Digital coherent receivers for long-reach optical access networks," *J. Lightw. Technol.*, vol. 31, no. 4, pp. 609–620, Feb. 2013.

[3] N. Suzuki, H. Miura, K. Matsuda, R. Matsumoto, and K. Motoshima, "100 gb/s to 1tb/s based coherent passive optical network technology," *J. Lightw. Technol.*, vol. 36, no. 8, pp. 1485–1491, Apr. 2018.

[4] R. M. Ferreira, A. Shahpari, J. D. Reis, and A. L. Teixeira, "Coherent UDWDM-PON with dual-polarization transceivers in real-time," *IEEE Photon. Technol. Lett.*, vol. 29, no. 11, pp. 909–912, Jun. 2017.

[5] D. Lavery, B. C. Thomsen, P. Bayvel, and S. J. Savory, "Reduced complexity equalization for coherent long-reach passive optical networks [invited]," *J. Opt. Commun. Netw.*, vol. 7, no. 1, pp. A16–A27, 2015.

[6] M. S. Faruk and K. Kikuchi, "Adaptive frequency-domain equalization in digital coherent optical receivers," *Opt. Exp.*, vol. 19, no. 13, pp. 12 789–12 798, 2011.

[7] J. Qi and F. N. Hauske, "Multiplier-free carrier phase estimation for low complexity carrier frequency and phase recovery," in *Proc. OFC/NFOEC*, Mar. 2012, pp. 1–3.

[8] K. Matsuda and N. Suzuki, "Hardware-efficient signal processing technologies for coherent PON systems," *J. Lightw. Technol.*, vol. 37, no. 6, pp. 1614–1620, Mar. 2019.

[9] Z. Dong, F. N. Khan, Q. Sui, K. Zhong, C. Lu, and A. P. T. Lau, "Optical performance monitoring: A review of current and future technologies," *J. Lightw. Technol.*, vol. 34, no. 2, pp. 525–543, Jan. 2016.

[10] D. A. Morero, M. A. Castrillón, A. Aguirre, M. R. Hueda, and O. E. Agazzi, "Design tradeoffs and challenges in practical coherent optical transceiver implementations," *J. Lightw. Technol.*, vol. 34, no. 1, pp. 121–136, Jan. 2016.

[11] L. Jiang *et al.*, "Fast and adaptive chromatic dispersion compensation scheme for digital coherent systems utilizing two-stage estimation," *Opt. Exp.*, vol. 23, no. 12, pp. 16 177–16 183, 2015.

[12] C. Xie, "Chromatic dispersion estimation for single-carrier coherent optical communications," *IEEE Photon. Technol. Lett.*, vol. 25, no. 10, pp. 992–995, May 2013.

[13] L. Jiang *et al.*, "Chromatic dispersion, nonlinear parameter, and modulation format monitoring based on Godard's error for coherent optical transmission systems," *IEEE Photon. J.*, vol. 10, no. 1, Feb. 2018, Art. no. 7900512.

[14] D. Zibar, M. Piels, R. Jones, and C. G. Scheffer, "Machine learning techniques in optical communication," *J. Lightw. Technol.*, vol. 34, no. 6, pp. 1442–1452, Mar. 2016.

[15] D. Wang *et al.*, "Cost-effective and data size-adaptive OPM at intermediated node using convolutional neural network-based image processor," *Opt. Exp.*, vol. 27, no. 7, pp. 9403–9419, 2019.

[16] S. Yan *et al.*, "Field trial of machine-learning-assisted and SDN-based optical network planning with network-scale monitoring database," in *Proc. Eur. Conf. Opt. Commun.*, 2017, pp. 1–3.

[17] J. Li, D. Wang, S. Li, M. Zhang, C. Song, and X. Chen, "Deep learning based adaptive sequential data augmentation technique for the optical network traffic synthesis," *Opt. Exp.*, vol. 27, no. 13, pp. 18 831–18 847, 2019.

[18] D. Wang *et al.*, "Machine learning-based multifunctional optical spectrum analysis technique," *IEEE Access*, vol. 7, pp. 19 726–19 737, 2019.

[19] Z. Wang *et al.*, "Failure prediction using machine learning and time series in optical network," *Opt. Exp.*, vol. 25, no. 16, pp. 18 553–18 565, 2017.

[20] F. N. Khan, C. Lu, and A. P. T. Lau, "Optical performance monitoring in fiber-optic networks enabled by machine learning techniques," in *Proc. Opt. Fiber Commun.*, 2018, pp. 1–3.

[21] A. L. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Commun. Surveys Tut.*, vol. 18, no. 2, pp. 1153–1176, Second Quarter 2016.

[22] X. Wu, J. A. Jargon, R. A. Skoog, L. Paraschis, and A. E. Willner, "Applications of artificial neural networks in optical performance monitoring," *J. Lightw. Technol.*, vol. 27, no. 16, pp. 3580–3589, Aug. 2009.

[23] V. Ribeiro, L. Costa, M. Lima, and A. L. J. Teixeira, "Optical performance monitoring using the novel parametric asynchronous eye diagram," *Opt. Exp.*, vol. 20, no. 9, pp. 9851–9861, 2012.

[24] T. S. R. Shen, K. Meng, A. P. T. Lau, and Z. Y. Dong, "Optical performance monitoring using artificial neural network trained with asynchronous amplitude histograms," *IEEE Photon. Technol. Lett.*, vol. 22, no. 22, pp. 1665–1667, Nov. 2010.

[25] F. N. Khan, T. S. R. Shen, Y. Zhou, A. P. T. Lau, and C. Lu, "Optical performance monitoring using artificial neural networks trained with empirical moments of asynchronously sampled signal amplitudes," *IEEE Photon. Technol. Lett.*, vol. 24, no. 12, pp. 982–984, Jun. 2012.

[26] X. Wu, J. A. Jargon, L. Paraschis, and A. E. Willner, "ANN-based optical performance monitoring of QPSK signals using parameters derived from balanced-detected asynchronous diagrams," *IEEE Photon. Technol. Lett.*, vol. 23, no. 4, pp. 248–250, Feb. 2011.

[27] J. A. Jargon, X. Wu, H. Y. Choi, Y. C. Chung, and A. E. Willner, "Optical performance monitoring of qpsk data channels by use of neural networks trained with parameters derived from asynchronous constellation diagrams," *Opt. Exp.*, vol. 18, no. 5, pp. 4931–4938, Mar. 2010.

[28] A. Caballero *et al.*, "Experimental demonstration of a cognitive quality of transmission estimator for optical communication systems," *Opt. Exp.*, vol. 20, no. 26, pp. B64–B70, 2012.

[29] Z. Ghahramani, "Probabilistic machine learning and artificial intelligence," *Nature*, vol. 521, no. 7553, pp. 452–459, 2015.

[30] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, 2015.