Interpretable and visualized SHAP-based equalizer with feature selection in IMDD system

Chenyu Wang (1), Xin Miao (2), Yu Xi (1), Meihua Bi (1,2), Longsheng Li (2), Weisheng Hu (2)
(1) *School of Communication Engineering, Hangzhou Dianzi University, Xiasha Gaojiangyuan 2rd Street, Hangzhou, Zhejiang province, 310038, China. E-mail: bmhua@hdu.edu.cn*
(2) *State Key Laboratory of Advanced Optical Communication System and Networks, Shanghai Jiao Tong University, 800 Dongchuan Rd, Shanghai, 200240, China. E-mail: wshu@sjtu.edu.cn*

Abstract: We firstly leveraged the SHAP-based method to visualize and analyze trained equalizers in IMDD-based short-reach system and manifested in bandwidth-limitation PAM-4 system. As a result, half of the features are reduced without deteriorating system performance.

1. Introduction

Recently, higher requirements for short-reach optical communication have been put forward as multiplying capacity while keeping a low cost. Therefore, the intensity modulation and direct detection (IM/DD) systems are still significant candidates, of which transmissions with limited bandwidth optoelectronic devices are mostly concerned at present [1]. However, for these high-speed systems based on limited-band transceivers, performances are often deteriorated seriously by the inter-symbol interference (ISI) and cumulative dispersion of optical fibers [2]. To compensate for those linear and nonlinear transmission impairments, several pre/post digital equalizers have been adopted, such as decision feedback equalizer (DFE), feed-forward equalizer (FFE), Volterra, and advanced machine learning algorithm like support vector machine (SVM), boosting-based algorithm, variants of neural network (NN) [3-5]. However, most of the previously proposed equalizers involve a trade-off between computational complexity and system performance. Linear equalizers such as FFE shows a limited performance even with a considerable tap number, while the nonlinear ones as Volterra are too complicated to employ in real scenarios. The machine learning (ML)/ NN -based ones [3-6] are demonstrated to achieve better results but demand too many features thereby also bringing high complexity. What's more, they are always restricted by the inexplicability of their trained models.

In this paper, we intend to solve the inexplicability problem of the ML-based equalizers for short-reach IM/DD systems. We firstly propose to apply the SHAP (SHapley Additive exPlanations) method [7] to improve the interpretability of the equalizers, and based on that we can reduce the complexity caused by redundant features. Here, SHAP is a mature framework which integrates a bunch of existing methods that can be used to explain the predictions of the black-box models. Specifically, by calculating its corresponding SHAP value, the original model of a ML/NN based equalizer can be represented by a linear model, which can be easier to visualize and analyze. Then, with the visualization interface of SHAP, the properties of each equalizer features can be learned, making the model interpretable. The proposed SHAP method is proved to be capable for a wide range of ML-based equalizers. Here in this work, we verify it in the 50Gb/s PAM-4 per wavelength IM/DD system with 10GHz optical devices, a simple SVM-based equalizer taken as an example. By calculating the global importance of each feature based on the SHAP value, half of the features used for training are reduced without deteriorating system performance. By analyzing the error cases in the test set, the feasibility of optimizing and constructing new powerful features for the equalizer is discussed at the end part.

2. Principle of the ML/NN-based equalizer with SHAP explainer

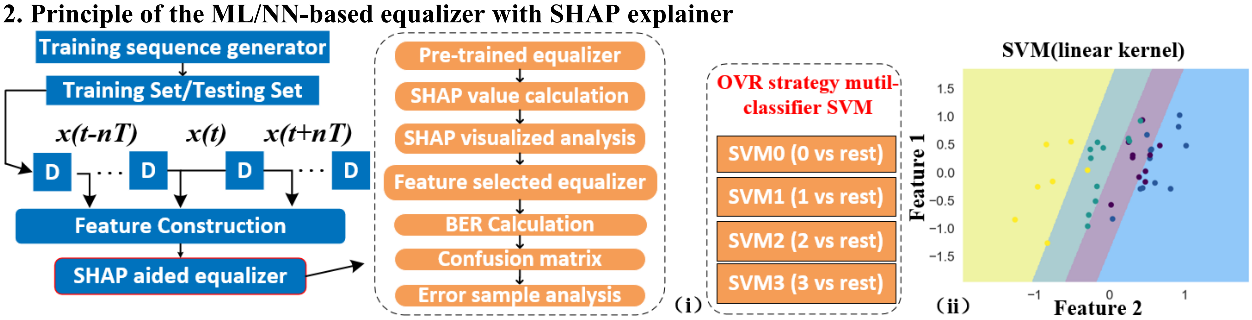

Fig.1. SVM-based equalizer with SHAP scheme Inserts: (i)OVR multi-classification strategy (ii) visualization for decision boundary of the 2-features SVM classifier (OVR)

In this section, the SVM-based equalizer combing with SHAP method is used to present our scheme. As shown in Fig. 1, the equalization processing is carried out in the receiver end of IM/DD system. The sequence produced by the training generator is used as labels and the corresponding received symbols as the original data, and then they are constructed as the training set and the test set, which are utilized to train and test the equalizer. Besides, as our previous work [8], through the delay module, the feature of received data can be conducted according to the channel characteristics. Based on these data, the SHAP aided SVM-based equalizer with 'One vs Rest'(OVR) classification strategy [9] can be performed in PAM-4 system. This specific equalizer is realized as shown in Fig.1. Here, the pre-trained equalizer is the common SVM-based equalizer, and then it is fed into the SHAP value calculation module. To make this equalizer more intuitive, the schematic of OVR classification strategy is presented in insert (i) of Fig.1 and the decision boundary of the OVR based SVM with the linear kernel is visualized in insert (ii) of Fig. 1. With the help of SHAP method, the contribution of each feature (Shapley value) for the prediction can be calculated, which can explain the equalization results of each sample instance in the original equalizer. As described in the reference [7], these Shapley values essentially can be regarded as the additive feature imputation method namely one linear model. Formally, this model can be specified as,

$$g(x_i) = \phi_0 + \sum_{j=1}^{M} \phi_j^i \tag{1}$$

where x_i is the *i-th* sample, *g(x)* is the explanation model, *M* is the dimension of the feature vector, and ϕ_j^i is the Sharpley value of feature attribution of the *j-th* feature in the *i-th* sample. With this equation, the explanation model *g(x)* can well match the original SVM-based equalizer model. Consequently, this linear model is used to explain the original equalizer and select the optimal feature sets. As presented in the flow chart of Fig.1, with the aid of the Sharpley value, a serious of visualizations of the equalizer can be performed in SHAP visualized analysis module. At the same time, the corresponding important features namely the effective features for equalizer can be selected. The global importance of each feature is calculated by $I_j = \frac{1}{n}\sum_{i=1}^{n}|\phi_j^i|$, where n is the number of samples used for training. Accordingly, the optimal features are extracted and then input to feature selected equalizer for realizing the final equalization and scheme testing. Owing to this operation, only a few features are employed in the equalizing process, hence reducing the complexity of the equalizer without any performance deterioration. With the trained equalizer, BER of test data is calculated. And, the confusion matrix is also visualized to further evaluate the performance of the feature-optimized equalizer. At last, these results are input into the error sample analysis module for further analyzing the error cases and the feasibility of optimizing or constructing new powerful features.

3. Experience setup and results discussion

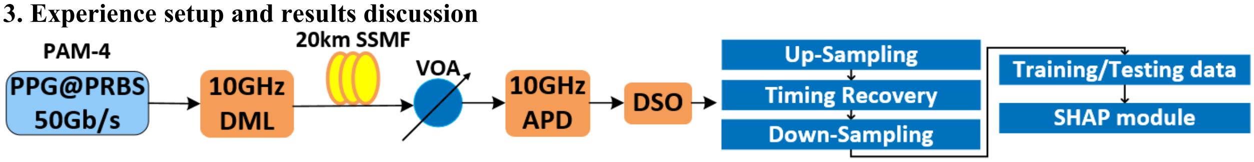

Fig. 2. Experimental setup of our scheme

To manifest our scheme, an IMDD system experiment is conducted as shown in Fig. 2. At the transmitter side, data with 2^{15}-1 pseudorandom binary sequence (PRBS-15) columns length are generated offline as the data source of the 50Gb/s PAM-4 signal. These mapped PAM4 signals and the random training sequence are then uploaded to a pulse pattern generator (PPG) to generate the transmitted signal which will then be utilized to drive a commercially available direct modulation laser (DML). This DML has ~18 GHz modulated bandwidth and 1311-nm working wavelength. The output power is measured to be 10dBm. After 20 km SSMF transmission, the optical signals reached a variable optical attenuator (VOA), which is employed for receiver sensitivity measurement. And then the signal is detected by an avalanche diode (APD) with a 3dB bandwidth as 7GHz at the receiver end. The detected signals are then captured by an 80-GSa/s sampling rate oscilloscope (DSO, LeCroy SDA845Zi-A) and the signal is sampled for the offline DSP module. For the DSP part, firstly, upsampling, clock recovery, and downsampling are conducted in MATLAB to obtain the raw structured tabular data that can be used for machine learning training. Then these data are used to construct the feature vector as shown in the feature construction module in Fig. 1. After that, we have all the data needed for the SHAP based equalization scheme which will then be implemented in the python environment.

As shown in Fig.3 (a), we give the SHAP summary plot for the decision of label 0. Before discussing the results, it is important to note that according to the mapping of Eq. (1), the sum of the Shapley values of each sample is numerically equal to the output of the learner, so we can quantify its effect on the sample prediction by the SHAP value of each feature. Obviously, the smaller the value of feature 18 is, i.e., the smaller the amplitude of the symbol is, the larger SHAP value the equalizer tends to give it. And this also means that in the SVM0 classifier and under the premise of correct classification, the samples with smaller feature18 values are further away from the SVM0 hyperplane, which means a higher probability of correct classification. Similarly, we can also analyze the properties of other features in equalization. To get a more intuitive perspective of the impact of each feature on the equalization of one symbol, we randomly select a sample with label0 that is correctly classified and present its decision plot in Fig.2 (b). To be concise, the model output value has been converted to the probability of a positive class (probability that the sample label is 0) [10]. What should be noticed is that, many features have almost no influence on the prediction result. Therefore, in Fig.3 (c), we give the global feature weights for each feature, and the features are sorted in descending order by weight. It can be intuitively seen that the contribution of some features is very low, meaning that there exsits many redundant or irrelevant features. Therefore, in Fig.4 (a), we give a comparison of the BER results for the original feature construction and feature selection (FS) scheme with the different number of features. It can be seen that, for the original 37-taps SVM scheme, the FS-SVM scheme can achieve the original result using only 18 features. To verify the validity of this feature selection scheme, we evaluate the BER performance in different received optical power for both B2B and 20km transmission. The results show that our scheme is consistently effective. To further evaluate the performance of the equalizer, we give the confusion matrix for the received optical power at -14dBm in Fig. 4. (c). It can be seen that classifiers tend to make more errors when predicting a sample with a true label of 2 as 3.

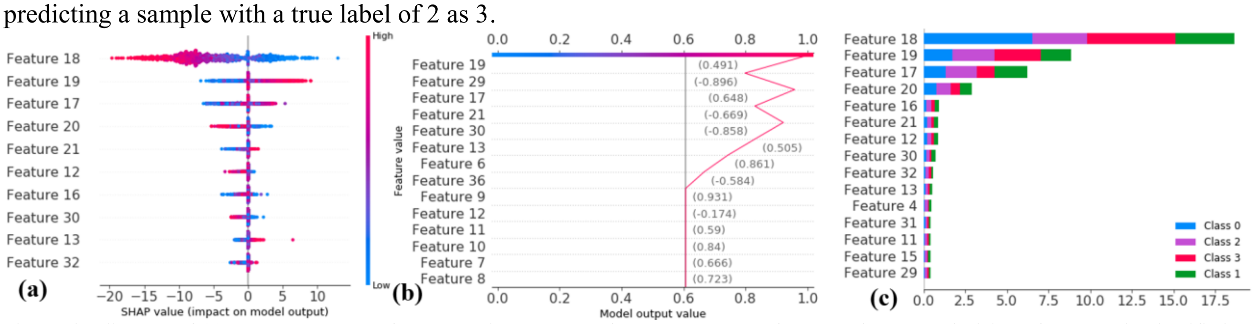

Fig.3. The diagram of (a) SHAP summary of SVM0(only present 10 features among 37 features) (b) SHAP decision of a correctly-classified 0-label sample (only present 15 features among 37 features) (c) SHAP global feature importance of each feature (only present 15 features among 37 features)

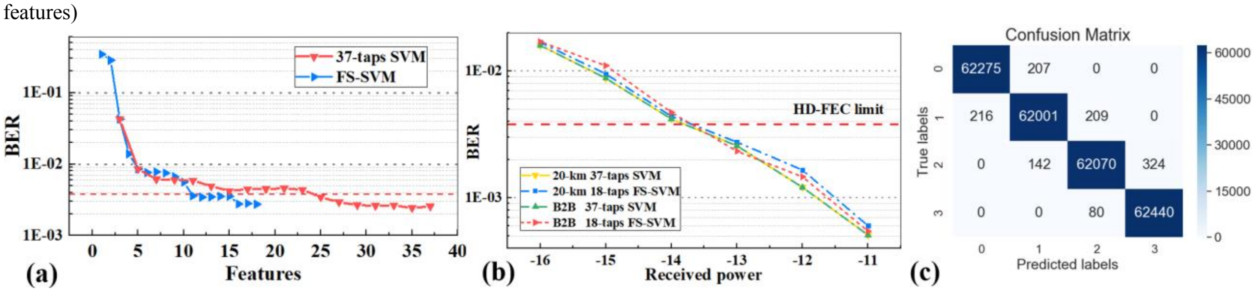

Fig. 4. The measured BER curves versus (a) different features used in 37-taps SVM (3000 training length) and 37-taps based feature selection SVM(3000 training length) scheme.(b) received optical power after 20km transmission and B2B case (c) Confusion matrix of 20-km FS-SVM scheme at -14dbm received optical power.

For these error points, we can use the SHAP decision plot to analyze. Besides, for the selected features we can further evaluate them by SHAP summary plot. Due to space limitations, only a brief analysis is given here, but with these visualizations and analyses, we assume that pre-equalization for the equalizer as well as qualitative and quantitative evaluation for higher-order features can be achieved.

4. Conclusion

The interpretable and visualized SHAP method used in ML/NN-based equalizer is proposed, and demonstrated in IMDD system with SVM equalization. By utilizing Shapley value, the optimal features are selected, which can decrease features without deteriorating performance. To verify its effectiveness and consistency, the experiment for PAM-4 signal in both B2B and 20km transmission are conducted. Results show the validity of our scheme. And, the confusion matrix @-14dBm received power is given, and discussed the analyzing the error cases to further optimize the features. Therefore, this SHAP-based method can be one of universal technique used in more ML/NN-based equalizers for optimizing the ML/NN-based equalizer and reducing their complexity and improving interpretability.

Acknowledgment

This work was jointly supported by the Natural Science Foundation of Zhejiang Province [No. LY20F050004] and the Scientific research project of Zhejiang Provincial Department of Education [No. Y202044258].

5. References

[1] S. Barthomeuf, et al, in OFC, 2019, W4J.4.
[2] L. Xue, et al. J. Light. Technol.38, 583–589 (2020.
[3] K.Zou, et al. in OFC, 2020, M1I-2.
[4] C.Wang, et al. in ACP,2020, M4A-325.
[5] L.Yi, et al. in OFC, 2020,T4D.3.7.
[6] T. Liao, et al. Opt. Lett.45, 5113–5116 (2020).
[7] S. M.Lundberg et al. in the 31st NIPS, 2017, 4768-4777.
[8] M. Bi, et al. Opt. Express 27, 25802–25813 (2019).
[9] A E. L.llwein, et al, JMLR, 113–141 (2000).
[10] J. Platt et al., Advances in large margin classifiers.10(3), 61–74 (1999).