A Versatile NN-Equalization for 50Gbps TDM PON Burst Uplink

Chenhui Ye, Xiaofeng Hu, Dongxu Zhang and Kaibin Zhang

Nokia Bell Labs, Shanghai, China (chenhui.a.ye@nokia-sbell.com)

Abstract: *To solve signal equalizer's vulnerability due to ONU uplink channel diversity in TDM PONs, we discovered that having a ONU-distinguisher Sub-NN as a steerer can effectively enhance the Main-NN's equalization resilience. We showcased such a commonly configured resilient NN serving 50Gb/s bursts from 6 ONUs.*

Introduction

In the quest for future PONs that can provide 50Gb/s+ access speed, digital signal processing (DSP) is about to play a crucial role in counteracting channel impairment and offloading the demand on high-end optics to cost-effective manners in software [1]. To name a few, feed forward equalizer (FFE), a simple but effective recipe for inter-symbol interference (ISI), has been widely adopted in high-speed serial communications. Lately, Volterra has been included into researchers' toolkit as nonlinearity emerges as a nonnegligible factor, making linear FFE fall out of line. Meanwhile, with the enormous success of machine learning in addressing implicit nonlinear problems these years, some preliminary types of neural network (NN) has been directly adopted for symbol recovery from channel distortion. Noticeable performance enhancement over traditional signal treatment has been constantly reported only with decent trainings based on unbiased samples [2-6].

Nevertheless, most of reported NN for signal equalization (NN4SE) schemes are targeted on point-to-point (P2P) scenarios where the end-to-end (E2E) link response is fixed and specific. PON-oriented NN4SE pioneers in [5] have warned on the vulnerability of NN4SE when it's scaled up from P2P to multi-point-to-point (MP2P) scenarios, typically like PON uplink. For ONU channel diversity's sake, the magic power of a NN in treating implicit channel impairments will degenerate considerably if NN's parameters are kept rigid while ONU uplink channels are reasonably diverse. By saying ONU diversity, it includes multi-aspect differentiations like in ONU's fiber distance, operating wavelength and components' bandwidth response, to name a few.

Increasing the NN depth and width in the hidden layers theoretically can enhance NN's capability in extracting channel features and augment NN's robustness against channel diversity, but the universality of NN equalizer against gain out of the DSP-complexity turns out to be not effective [5]. Therefore, some other NN4SE strategies are needed not only for nonlinearity management but also allowing OLT with universality in treating diverse ONUs.

In this paper, we'll introduce a resilient NN that can automatically steer its equalization degree for each uplink channel according to some implicit feature machine learnt by itself, and therefore no NN weights reconfiguring is needed. It's done by the introduction of a Sub-NN that serves as a ONU distinguisher embedded inside the Main-NN and has its outputs remerged to the intermediate hidden layers of the Main-NN to manipulate the equalization function subtly. In the proof-of-concept tests of 50Gb/s-NRZ uplink over 20GHz-class C-band direct modulation laser (DML), 3 far (20km) ONUs with variant-BW DML and another 3 near (3km) ONUs all can meet <6E-3 BER threshold before FEC. Further investigation and interpretation on the ONU-distinguisher Sub-NN indicates that universal-NN4SE for PON uplink is not only robust but also can be very low-complex.

Universal-NN4SE w/ ONU-distinguisher

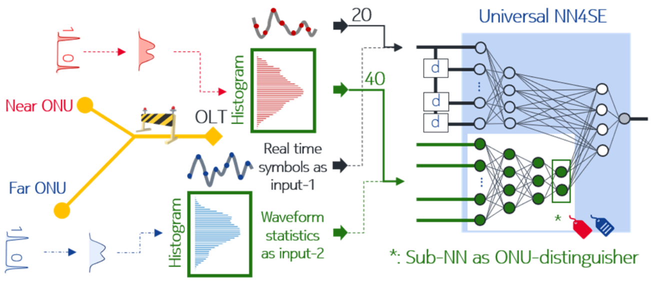

Fig.1 OLT side universal-NN4SE architecture

The secret recipe of the Universal-NN4SE lies in the specialized NN structure with a Sub-NN embedded inside a Main-NN as depicted in **Fig.1**. Besides a regular input-1 for real time symbols, there's an input-2 for waveform statistics in form of amplitude-histogram. Input-1 contains 10 historical, a current and 9 posteriori, in total 20 received symbols to allow Main-NN to learn the mysterious ISI mechanism. The histogram, input-2 to Sub-NN, is believed to have been enciphered with the footprint of a E2E uplink channel response which is implicit to be described. The Sub-NN is designed as such for the purposes of 2-fold, 1st to decipher the histogram and extract some low dimensional ONU-codes, 2nd to have the codes merged back to the later part of Main-NN to carefully steer its equalization degree to that ONU.

For training, the Universal-NN4SE needs to be fed with received & targeted symbol pairs together with corresponding statistical histogram from all ONUs. The embedded Sub-NN gets trained simultaneously and in synergy with the Main-NN, meaning the ONU-codes learnt by Sub-NN is not subjectively assigned, but solely determined by its code-width and the way it's connected to the Main-NN.

Experimental Setup and Results Discussion

To emulate the stacked challenges of 'ONU-diversity over nonlinear ISI' in PON uplink, we setup a test bed as shown in **Fig.2**. 50Gb/s NRZ data are pre-filtered in AWG (Keysight 8194A) by 6 different sets of low-pass FIR filter to mimic bandwidth (BW) differences among DMLs. The filtered signals are then modulated by a common C-band DML with a 10dB-BW of ~18.9GHz. A segment of 20km fiber and another of 3km are used to introduce the second dimensional diversity in terms of fiber length. In receiver side, OLT, optical amplification and a 33GHz (@ 3dB-BW) PIN is used for signal detection before quantizing in a 160GSa/s ADC (Keysight Z592A).

Before any DSP, the after-propagation signals from 6 ONUs are plotted in histogram as the inset-(a) in **Fig.2** shows, with 40 bins for amplitude statistics. According to one's domain-experience, high-level comments can be given on some 'hidden messages' sealed beneath the histograms, a 1-D eye-diagram: e.g.,

1. Semi-duobinary pattern can be observed in Histogram-4/5/6 mainly due to bandwidth-restriction;
2. Asymmetric eye-skew can be noticed in Histogram-1/2/3 due to chirp-dispersion interaction.
3. Apparently, the 6 histograms from far (1/2/3) and near (4/5/6) groups of ONUs can be broadly classified into 2 main clusters simply by their fiber length, indicating that the impact from fiber length difference is the dominating factor over the DML bandwidth on C-band.

However, there are still many feature similarities and differentiations beyond explicit description in language. Above all, what really matters is if/how the machine can spot these major and minor differentiations, and more crucially is if/how to translate them into machine language to steer the equalization degree appropriately.

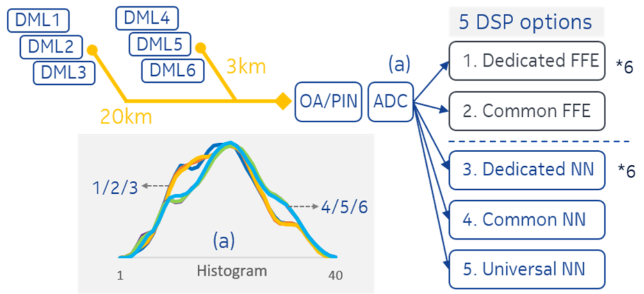

Fig.2 Experimental setup with (a) histograms

To evaluate the effectiveness of the proposed Universal-NN4SE, we compare in total 5 DSP approaches in parallel, namely 1. Dedicated linear FFE (*6), 2. Common linear FFE, 3. Dedicated NN (*6), 4. Common NN and 5. Universal-NN. By saying dedicated, it means there are 6 individual NN/FFEs (DSP-1&3) trained and tested exclusively for each of 6 ONUs; common (DSP-2&4) means have a public NN/FFE trained up by samples scrambled amongst 6 ONUs and then be tested separately for each ONU; similar to the common case, {in/ out} data pairs in the form of {received signals + histogram/ targeted symbol} from 6 ONUs are used for the Universal-NN training (DSP-5). Mersenne twister pseudo-random number sequence [6] of 100k is used for training, another 60k is used for test. Both FFEs in DSP-1&2 are structured in 20*1 with a linear activation function (AF); both NNs in DSP-3&4 are structured in 20*10*4*1 with tanh typed nonlinear AF. As for the Universal-NN in DSP-5, the Main-NN is same as 20*10*4*1 and the Sub-NN is structured as 40*10*4*C where the C represents the output width, Sub-NN's outputs are merged into the last hidden layer as in **Fig.1**.

Detailed evaluation on 5 DSP approaches in terms of BER vs receiver optical power are plotted in **Fig.3**. The curve cluster of DSP-1 shows that traditional FFE can take pretty good care of channel impairments for near ONUs (4/5/6) but is underqualified for far ONUs (1/2/3). The curve clusters of DSP-2 indicate that performances get worsened (by different degrees) when the dedicated FFE is substituted by a common FFE configuration. Obviously the common FFE cannot converge well under wiggling targets in common training. On the contrary to linear Dedicated-FFE in DSP-1, Dedicated-NN in DSP-3 can satisfy both near and far ONUs by having the hybrid linear and nonlinear channel issues well managed. Similarly, the magic of NN degenerates considerably when the Dedicated-NN (DSP-3) is replaced by a Common-NN (DSP-4) which has also been reported in [5]. On the contrary, the Universal-NN (DSP-5) provides quite consistent efficiencies to all near and far ONUs, which means the Sub-NN is indeed functioning like an ONU-distinguisher as we designed it for.

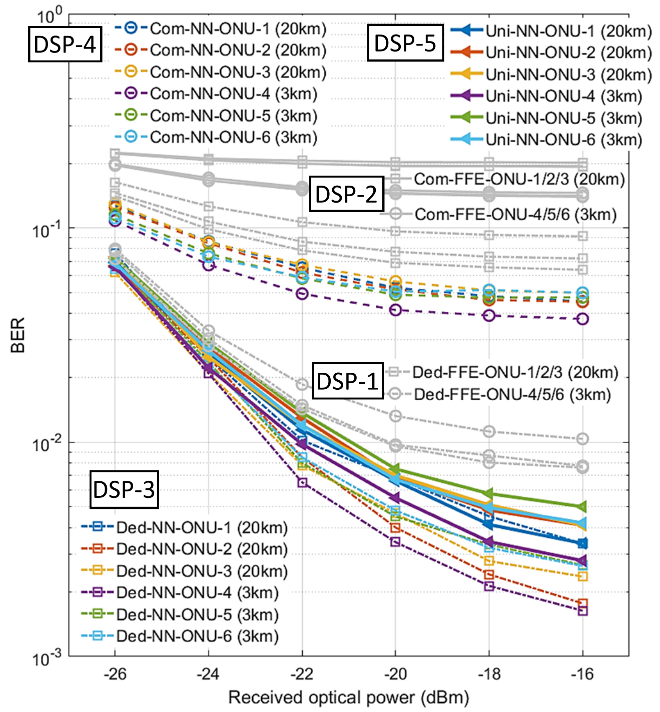

Fig. 3 Experimental results and comparisons

Deeper Investigation on Sub-NN

It's not surprising for a machine to learn classification, but it's not intuitive what the Sub-NN has learnt for its output is not restricted during the Universal-NN4SE training. Above all, to demystify the underlying mechanism, especially how such 'unconsciously' learnt knowledge in Sub-NN steers the equalization degree, is decisive in practical Universal-NN4SE designs.

First, Sub-NN needs to be detached from the main body to see the intermediate variables, i.e., the ONU-codes. A self-contained NN is cloned from the Sub-NN {with 40*10*4*C NN-structure} and is fed with the 6 histograms to observe the ONU-codes. **Fig.4-(a)** shows 6 ONU-codes in a 4-D radar chart form when C = 4. It matches pretty well with our previous domain-intuitive qualitative comments (in Section 3) that ONUs with the same fiber length have similar ONU-code value patterns. In fact, the ONU-codes learnt by Sub-NN each time can be totally different like (b) and (c), however, it hardly impacts the final equalization accuracy of the whole Universal-NN4SE (DSP-5) for all 6 ONUs shown in **Fig.3**. It implies that there is one global minimum in the overall optimization space but with multiple optimization minima in the Sub-NN space. In other words, there's dimensional redundancy in Sub-NN's output ONU-code width. It's been self-proved by the clustering of the values on the 4th-axis in (b) amongst all 6 ONUs.

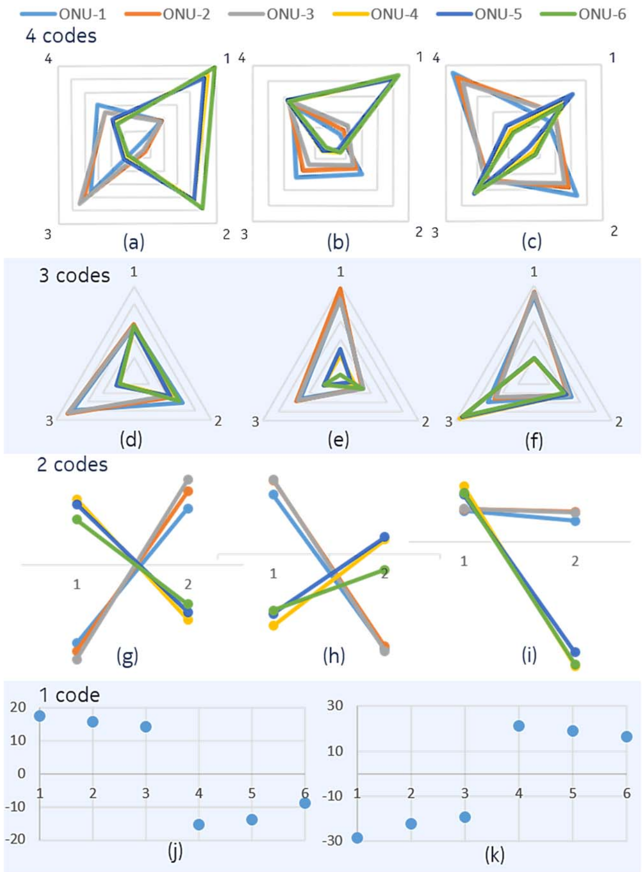

Fig.4 BER curve vs RL iterations

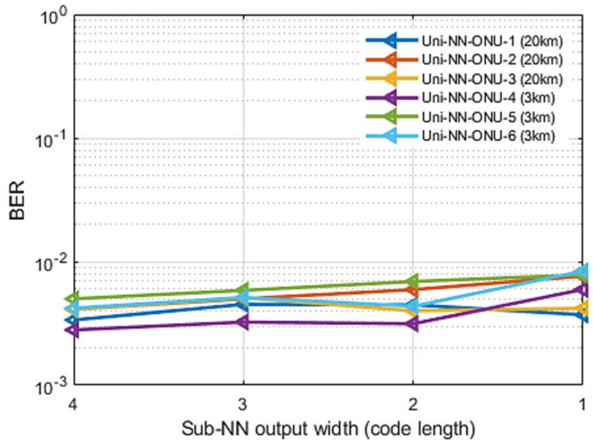

Fig.5 BER vs Sub-NN output width (ONU-code-length)

Therefore, we gradually reduce Sub-NN's output-width, while having the rest of NN structure remained to investigate the impact from the ONU-code width. **Fig.4** Row-2/3/4 shows different ONU-code value patterns when Sub-NN's output-width is set to 3, 2, and 1, respectively. Similar to Row-1, traces that indicate code-width redundancy can also be observed in 3-codes cases like 1st-axis in (d) & 2nd-axis in (f), even slightly observed in the 2-codes cases like 1st-axis in (i). Recalling that the elaborately controlled diversity in experiments are precisely 2 independent dimensions, i.e., BW and fiber length, the redundancy becomes explainable after all, until ONU-code is further reduced to 1. (j) & (k) in **Fig.4** are the only 2 possible ONU-code patterns observed in a serial of tests, it's interesting to see that the 2-D diversities in between ONU-1/2/3 and ONU-4/5/6 have been projected to 1-D by the Sub-NN being restricted in single value code width.

BER vs ONU-code-length shown in **Fig.5** indicates some but slight degradation in performance can be seen in the code width reduction in our experimental test scenarios. But it doesn't imply that single variable ONU-code is enough for all PON systems, caution is needed to refer to the ONU dynamic in a given PON system.

Considering the stability of the distribution network of a PON, Sub-NN can be boiled down to its ONU-code itself (framed in **Fig.1**) once the whole NN has been trained up. Therefore, the DSP-complexity of the Universal-NN4SE will be very comparable to a regular Dedicated-NN.

Conclusion

A ONU-distinguisher Sub-NN is newly introduced in OLT and embedded inside the Main-NN to help elaborately steer the equalization according to the ONU-codes it learnt from statistical waveform features for each ONU in PON uplink. Armed with such Universal-NN4SE, OLT can well mitigate linear/nonlinear channel impairments for variant ONUs with a common NN configuration without frequent NN-weights reconfiguration every burst. Further research indicates that the delta DSP-complexity of such approach will be negligible over a regular NN that allows 20G-DMLfor 50Gb/s uplink services.

Acknowledge

Special acknowledge to Jochen Maes in Bell Labs Antwerp, Belgium for inspiring discussion and inputs.

References

[1] Anet Neto L, et al. “Considerations on the Use of Digital Signal Processing in Future Optical Access Networks.” JLT, 2020.

[2] Estaran J, et al. "Artificial neural networks for linear and non-linear impairment mitigation in high-baudrate IM/DD systems." ECOC2016.

[3] Chagnon M, et al. “Experimental demonstration of a dispersion tolerant end-to-end deep learning-based IM-DD transmission system” ECOC2018.

[4] Ye C, et al. "Recurrent neural network (RNN) based end-to-end nonlinear management for symmetrical 50Gbps NRZ PON with 29dB+ loss budget." ECOC2018.

[5] Houtsma V, et al, “92 and 50 Gbps TDM-PON using neural network enabled receiver equalization specialized for PON.” OFC2019.

[6] Chuang C Y, et al, “Study of Training Patterns for Employing Deep Neural Networks in Optical Communication Systems”, ECOC2018.