Clustering G-PON Field Data to Improve Flexibility in Next Generation PON Systems

Gaël Simon(1), Philippe Chanclou(1), Fabienne Saliou(1), Jérémy Potet(1,2), Minqi Wang(1,3)

(1) Orange, 2 avenue Pierre Marzin 22307 LANNION, France, gael.simon@orange.com
(2) Université Rennes, CNRS, Institut FOTON - UMR 6082, F-22305 Lannion, France
(3) IMT Atlantique, 665 Avenue du Technopôle, 29200 Plouzané, France

Abstract *We study two scenarios for more flexibility in next generation PONs, with the help of the collected data from 4 million deployed ONUs. Clustering the ONUs over the ODNs shows that a clustering based on the OLT-ONU distance is the best solution for balanced clusters.*

Introduction

While Gigabit Passive Optical Network (G-PON) is nowadays widely deployed, discussions are ongoing about next PON generations[1-4]. International Telecommunication Union (ITU-T) started to release a set of standards on its 50Gb/s (downstream) solution called Higher Speed PON (HS-PON). The next generation PON will need to keep paying off the huge investment realized to deploy PON's millions of kilometers of fiber, while the cost of the new hardware should remain low. The high baud rate, coupled with the diversity of the transmission characteristics, increases the combinations to be covered: diversity of optical path distances for a given Optical Distribution Network (ODN) or between ODNs, diversity of Optical Path Losses (OPLs), diversity of transmitter/receiver (Tx/Rx) characteristics (bandwidth, gain, noise,...), diversity of optical fiber chromatic dispersion characteristics,...

HS-PON natively enables the use of Digital Signal Processing (DSP) to overcome optical transmission limitations. However, the point-to-multipoint topology of PON complicates the problems of finding a set of DSP parameters matching all possible transmission characteristics, and adds another diversity parameter to the already complicated problem[5,6]. We proposed in [7] to exploit the fact that the distances of the Optical Network Units (ONUs) of a given ODN are generally at the same distance from the Optical Line Terminal (OLT). The resulting optical path transmission characteristics are then locally simplified, also theoritically simplifying the parametrization of the DSP. Other works[8] proposed to increase the degree of flexibility in next generation PONs in exploiting ONU grouping through flexible modulation, flexible Forward Error Correction (FEC), ...

We propose here to exploit the field data that G-PON operators have collected, in order to qualify the already deployed infrastructure in the context of a HS-PON with flexible capabilities.

Two scenarios

We propose to assess two scenarios (Fig1):

- In the first, "ONU-based flexibility" (Fig1.a), each ONU is independently managed. In the case of FEC, the natural solution would consist in applying the FEC profiles according to the received power at the OLT side for upstream (US) or ONU side for downstream (DS). Note that according to XGS-PON recommendations[9], "in the downstream direction, FEC is statically configurable as on for all ONUs; in the upstream direction, the use of FEC is under dynamic control by the OLT". Also, the FEC can generally only be enabled or disabled, without specifying the FEC type, while the XGS-PON recommends[9] the use of Reed-Solomon [255, 239] or [255, 223]. Using a lower FEC overhead, the operator enables the "low loss optical paths" and the related customers to benefit from a slightly higher user bitrate, while making the management more complex for the operator. The corresponding scenario in terms of DSP management and its limitations is discussed in [7]. Such solution could permit to meet the technical limitation of DSP in HS-PON[10,11].
- In the second scenario, "Flexibility based on

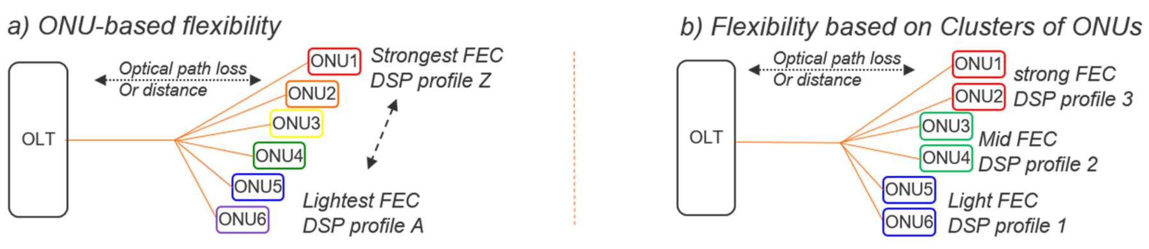

Fig. 1: The two flexibility scenario studied: "ONU-based flexibility" (a), "Flexibility based on clusters of ONUs" (b)

978-0-7381-4679-9/21/$31.00 ©2021 IEEE

clusters of ONUs” (see Fig1.b), the ONUs of a given ODN are clustered. A DSP profile is applied to each cluster, and so does the FEC profiles for US and DS. However, one of the problems is to correctly determine the optimal number of clusters and their characteristics.

Data exploitation considerations

Operators collect every day huge quantities of data from their infrastructure. We propose here to exploit the data from more than 4 million ONUs from Orange’s customers, in France. The received power at OLT and ONU side, but also the OLT-ONU distance, are already collected on a regular basis. The received power is provided by Received Signal Strength Indicator (RSSI) measurements with a 10^{-2}dBm resolution. The accuracy is limited to +/-3dB or +/-2dB for the ONU and OLT received power, respectively according to the recommendations. The measured distance resolution is 1m or 100m, depending on the OLT vendor’s implementation. Besides, the data used here are from the deployment in France, which mainly started in dense areas (then short distances) and it is still ongoing. The data distribution may look different in a few years.

ONU-based flexibility: results & discussion

The “ONU-based flexibility” (Fig. 1.a) classification study can be performed using directly the received power distribution, independently from the ODN affiliation. The black dash-dotted curve of Fig. 2 represents the statistical distribution of all received powers at OLT side of 4 million optical paths, while the blue dotted curve shows the corresponding cumulative distribution. We propose to discuss an example with 4 thresholds (-28dBm, -23dBm, -18dBm, -13dBm), aligned with the minimum sensitivity of -28dBm for the N1 Optical Budget class[9]. The five corresponding sub-classes (-Inf.⇔-28dBm,-28dBm⇔-23dBm, etc.) present the following densities: 3%, 22%, 73%, 2%, ~0%.

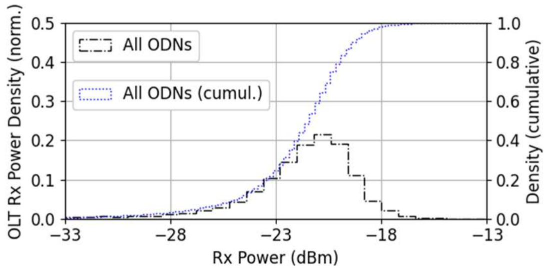

Fig. 2: Distribution of upstream Rx power

A 5dB linear step is totally arbitrary. Determining the optimal number of classes requires to evaluate the balance between the advantages of more flexibility, and the cost of a more complex hardware implementation, that allows all OLTs and ONUs of the field to manage the different classes. For example, the integration of 5 classes of fast processing algorithms enabling several levels of FEC would probably not be at a negligible cost, as would be the electrical power consumption. Also, neighbours customers could benefit from a totally different quality of service, depending on the maximum quality of their (optical) path as it used to be in xDSL. The latter could simply depends on a dirty connector. Such evolution could be seen as regression, with “a pay as the throughput you get” approach.

Flexibility based on clusters of ONUs: results & discussion

The second scenario, “flexibility on cluster basis” on Fig. 1, consists in grouping data having similarities, and apply specific parameters to match each of them, such as DSP parameters or FEC profiles. Cluster analysis regroup several families of algorithms and methods, such as hierarchical clustering or k-means[12,13]. We propose to use the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. The latter takes only two main parameters. The first is the minimum number of points required to form a cluster, which we set to one, since no point corresponding to a customer

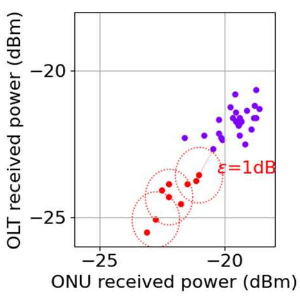

Fig. 3: OLT vs. ONU received power of ODN Alpha (2 clusters, ε=1dB)

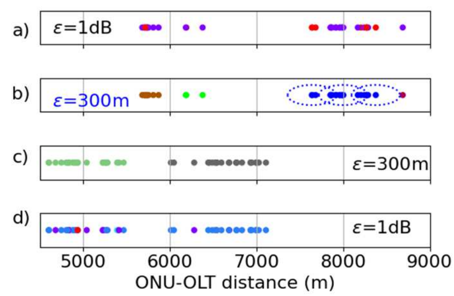

Fig. 4: Path distance of (a) ODN Alpha with Rx power clustering; (b): same than (a) with distance based clustering. (c) and (d): similar to (a) and (b), but with ODN Beta (ε=300m)

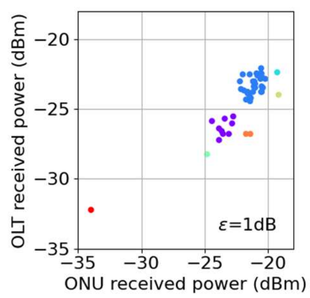

Fig. 5: OLT vs. ONU received power of ODN Beta (7 clusters, ε=1dB)

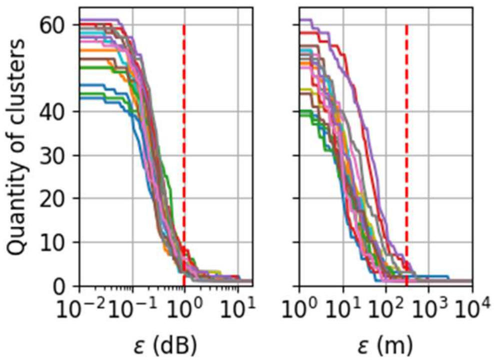

Fig. 6: quantity of clusters versus ε for Rx power clustering (a) and distance clustering (b)

can be left apart, even if it may look standing apart from other points. The other parameter, designed by ε, is the maximum magnitude between two samples, for one to be considered as in the neighbourhood of the other.

Fig. 3 shows an example of clustering with DBSCAN, using received powers of all ODN paths called "Alpha", with ε=1dB. In this example, two clusters exists (red and purple), with 9 and 27 ONUs respectively. As DBSCAN algorithm does not take into account the data units, we need to separate the data in two sets: clustering over OLT and ONU received power of the ODN ("Rx power" clustering), and clustering over ONU-OLT path length ("distance" clustering).

The ε value impacts deeply the number of clusters. Fig. 6 shows the number of clusters versus ε for Rx power clustering (a) and distance clustering (b) for 18 random ODNs with more than 40 ONUs each. When ε is low (ε<0.01dB or ε<1m), ε eventually gets lower than the data resolution, and the number of clusters meets the number of ONUs of the ODN (when no value duplication). On the opposite, when ε is high enough (ε>10dB or ε>20km), all data are included in the same cluster. With ε=1dB or ε=100-300m, the number of clusters for a given ODN should in most cases then be lower than 10, a "realistic" number of clusters, compared to "ONU-based flexibility". The latter can be seen also as the result of clustering with a high number of groups.

Fig. 4.a&b show also ODN Alpha, but with Rx power clustering (a) and distance clustering (b). Comparing both figures, one can see that the two methods produce different clusters, as Fig. 4.b shows 4 clusters with ε=300m. Besides, in case of ODN Alpha and ε=300m, a cluster made of a single ONU appears (Fig. 4.b, at ~8700m).

We propose to call σ the parameter describing the balance of the clusters, and define it as the size of the smallest cluster divided by the largest cluster of an ODN. σ~0 for unbalanced ODNs where at least one cluster is made of a single ONU, and σ=1 for uniformly distributed ODNs

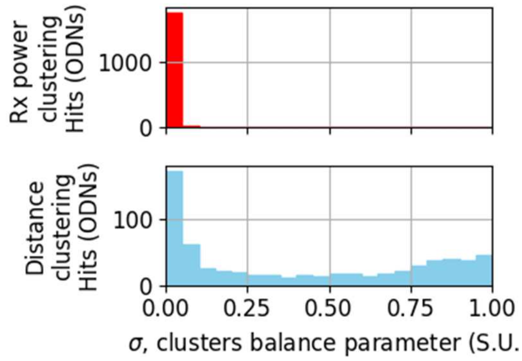

Fig. 7: Clusters balance for Rx power clustering (top) with ε=1dB and distance clustering with ε=300m (bottom)

where all clusters have the same size. For ODN Alpha (Fig. 3), σ=9/27=0.33 when ε=1dB.

Fig. 7 shows the distribution of σ for Rx power clustering (top) with ε=1dB and distance clustering with ε=300m (bottom) for respectively 1785 and 647 ODNs with more than 34 ONUs. ODNs with a single cluster are removed (29 and 1157 respectively), since it implies σ=1. One can see that in the case of Rx power clustering, σ is generally close to zero, meaning that in many cases clusters are made of a single ONU. This is for example the case of ODN "Beta" (σ=0.02), presented on Fig. 5. Clustering the ODNs in an Rx power fashion would induce clusters dedicated to isolated ONUs. On the opposite, distance clustering shows (Fig. 7.b) a more uniform distribution, meaning that distance clustering would unfortunately still bring isolated ONUs, but also balanced clusters on the ODN.

Conclusions

We studied two scenarios to improve the flexibility of next generation PONs in terms for example of FEC diversity or optical path transmission characteristic and DSP diversity. "Flexibility on ONU basis" is the most complete solution, but may lead to a difficult marketing in the sense that customers may benefit from different quality of experience. The hardware feasibility is also to be demonstrated. We also proposed flexibility on a cluster basis, consisting in grouping similar ONUs of the same ODN. The exploitation of G-PON field data using DBSCAN algorithm showed that ε=1dB or ε=100-300m lead to a realistic number of clusters (<10). We also showed that Rx power based cluster are not the same than distance clusters. Rx power based cluster lead to unbalanced clusters with isolated ONUs. Distance clusters are more balanced, even if isolated ONUs still exist.

Acknowledgements

This work was carried out in 5G-COMPLETE H2020-ICT-2019 and MARSAL H2020-ICT-52 projects.

References

[1] G.9804.series of recommandations, ITU-T, April 2021

[2] Jeremy Potet, Fabienne Saliou, Gaël Simon, Mathilde Gay, Laurent Bramerie, Philippe Chanclou, Monique Thual, « Assessment of an SOA for Burst Mode Pre-Amplification in Higher Speed PON hitting 30dB of Optical Budget", Optical Fiber Conference, 9th of June 2021 , OFC 2021, San Franciso, CA, USA

[3] B. Li et al., "DSP enabled next generation 50G TDM-PON," in IEEE/OSA Journal of Optical Communications and Networking, vol. 12, no. 9, pp. D1-D8, September 2020, doi: 10.1364/JOCN.391904.

[4] V. Houtsma and D. van Veen, "Optical Strategies for Economical Next Generation 50 and 100G PON," 2019 Optical Fiber Communications Conference and Exhibition (OFC), 2019, pp. 1-3.

[5] N. Kaneda, D. van Veen, A. Mahadevan and V. Houtsma, "DSP for 50G/100G Hybrid Modulated TDM-PON," 2020 European Conference on Optical Communications (ECOC), 2020, pp. 1-4, doi: 10.1109/ECOC48923.2020.9333248.

[6] C. Ye, X. Hu, D. Zhang and K. Zhang, "A Versatile NN-Equalization for 50Gbps TDM PON Burst Uplink," 2020 European Conference on Optical Communications (ECOC), 2020, pp. 1-3, doi: 10.1109/ECOC48923.2020.9333360.

[7] G. Simon, F. Saliou, P. Chanclou, L. A. Neto and H. Hallak Elwan, "50Gb/s TDM PON Digital Signal Processing Challenges: Mining current G-PON Field Data to Assist Higher Speed PON," 2020 European Conference on Optical Communications (ECOC), 2020, pp. 1-4, doi: 10.1109/ECOC48923.2020.9333167.

[8] R. Borkowski et al., "World's First Field Trial of 100 Gbit/s Flexible PON (FLCS-PON)," 2020 European Conference on Optical Communications (ECOC), 2020, pp. 1-4, doi: 10.1109/ECOC48923.2020.9333413.

[9] G.9807.1 : 10-Gigabit-capable symmetric passive optical network (XGS-PON), ITU-T, https://www.itu.int/rec/T-REC-G.9807.1-201606-I/fr

[10] S. Barthomeuf et al., "Equalization and Interoperability Challenges in Next Generation Passive Optical Networks," 2019 Optical Fiber Communications Conference and Exhibition (OFC), 2019, pp. 1-3.

[11] Hamza Hallak Elwan, Fabienne Saliou, Gael Simon, Luiz Anet Neto, and Philippe Chanclou, "Towards Unified Channel Equalization to Meet Transmission Requirements for Next Generation Passive Optical Networks", EuCNC, 2021

[12] Géron, A. (2017). Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems. Sebastopol, CA: O'Reilly Media. ISBN: 978-1491962299

[13] Tibshirani, R., Walther, G. and Hastie, T. (2001), Estimating the number of clusters in a data set via the gap statistic. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63: 411-423. https://doi.org/10.1111/1467-9868.00293