Identification of optical links with heterogenous fiber types in a production network

Emmanuel Seve, Sebastien Bigo and Patricia Layec
Nokia Bell Labs, 12 rue Jean Bart, 91300 Massy, France.
emmanuel.seve@nokia-bell-labs.com

Abstract: We develop a technique to identify fiber type within heterogeneous network links using correlation between lightpath accumulated dispersions. We successfully identified fiber types from real data issued from a continental-size production network running live.

1. Introduction

The incomplete knowledge of network topology and of exact performance of deployed hardware forces operators to increase design margins, and therefore to prevent the full use of their resources [1-2]. In this paper, we focus on the inaccuracies of fiber types, stemming from e.g., poor inventory or splicing mistakes. Methods based on fiber-longitudinal monitoring [3-5] have been recently proposed against these inaccuracies but require transferring a large quantity of information to the controller card and therefore cannot be implemented with the current states of technologies. In [6], we proposed a Mixed Integer Linear Programming (MILP) algorithm to automatically discover the fiber type of homogeneous links (same type of fiber between network nodes) by correlating the accumulated dispersion (CD) of all established network lightpaths measured by coherent receivers. Such CD information is already monitored and transferred to the controller card in installed recent commercial coherent receivers.

In this paper, we generalize the fiber type identification to heterogenous links (various type of fiber between network nodes) thanks to a multi-solutions ranking. We also successfully demonstrate successful automatic identification with real CD measurements from a production network running live.

2. Method

Our method is performed in two steps: extract likely heterogeneous links in the network (step 1) and then, identify fiber types within these heterogeneous links (step 2). The search areas for the unknown values of the averaged CD and CD slope of each link or fiber span is shown in Figure 1 for the seven most common fiber types, accounting for production ranges of the fiber industry. The white crosses stand for the so-called "typical" values for each fiber type. Our algorithm in [6] prioritizes for the identification of fiber types when there is a single type between nodes and can report and rank every set of fiber types (called "solution" in the rest of the paper) meeting this condition as well as all MILP constraints inside the fiber search areas. If it comes across one or several heterogeneous link (as shown for instance by the black bullet in Fig. 1), it reports no solution since the space outside fiber ranges is not explored.

2.1 Step 1: Determine the nature of the network links (homogeneous or heterogeneous).

We propose a generalized method to deal with heterogeneous links. The main idea is to allow the algorithm to automatically put aside the fiber type constraints and explore the complete space in terms of dispersion and dispersion

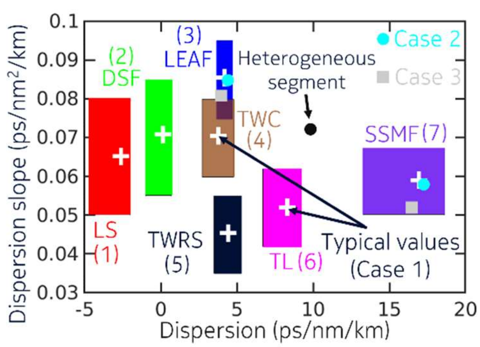

Fig. 1. Fiber characteristics (dispersion and dispersion at 1550nm). White crosses correspond to the typical values. Cases 1-3: fiber characteristics of the heterogeneous link 1 in the tested topology (see section 3).

Table 1: Step 2 MILP equations (generalized method)

$$\sum_{k=1}^{N} CD_k^S = CD^L \ , \ \sum_{k=1}^{N} CD'^S_k = CD'^L \quad (1)$$

$$min\left(\sum_{k=1}^{N}\left[\left|CD_k^S - CD_k^{typ}\right| + \left|CD'^S_k - CD'^{typ}_k\right|\right]\right) \quad (2)$$

$$CD_k^{typ} = \sum_{j=1}^{N_f} CD_j^{typ} FT_{j,k}, CD'^{typ}_k = \sum_{j=1}^{N_f} CD'^{typ}_j FT_{j,k} \quad (3)$$

$$P_i = \prod_{k=1}^{N} \frac{1}{\sigma_k\sqrt{2}} e^{-\left(\frac{CD_k - CD_k^{typ}}{\sigma_k}\right)^2} \prod_{k=1}^{N} \frac{1}{\sigma'_k\sqrt{2}} e^{-\left(\frac{CD'_k - CD'^{typ}_k}{\sigma'_k}\right)^2} \quad (4)$$

$$R_i = P_i / \sum_{i=1}^{N_{sol}} P_i \quad (5)$$

$$\left(1 + \frac{{\sigma_N}^2}{{\sigma_i}^2}\right)\left(\frac{CD_i - CD_i^{typ}}{\sigma_i}\right) + \dots$$
$$\sum_{j\neq i}^{N}\left(\frac{CD_j - CD^{typ}}{\sigma_j}\right) = CD^L - \sum_{i=1}^{N} CD_i^{typ} \quad (6)$$

slope only when it is needed to get a non-null solution. We introduced binary output variables θ_L^i to control the presence of the constraints related to the CD and CD slope of the link (Eqs (3-6) in our previous work [6]). These constraints are automatically removed from the complete set of constraints when output variables θ_L^i are equal to 0, corresponding to heterogeneous links. To prevent the algorithm to affect null values for all links and considering that all network links are heterogeneous, we introduce a penalty in the objective function for each null value of θ_L^i, pushing the algorithm to find solutions with the minimal number of heterogeneous links. The objective function which is now equal to $\sum_{j=1}^{N^{LP}} \Delta CD_j$ - $\sum_{i=1}^{N^L} \theta_L^i$ where ΔCD_j is the CD measurement accuracy for the lightpath j. $N^{LP}$is the number of lightpaths and $N^L$is the number of links in the network. We also propose to add a new constraint to in the MILP to impose a maximal number of different fiber types inside one link.

2.2 Step 2: Identify the (most probable) fiber types within the heterogeneous link.

After step 1, all heterogeneous links are identified ($\theta_L^i = 0$) and characterized by their accumulated dispersion CD^L and dispersion slope CD'^L. A new MILP (see Table 1, eqs. 1-3) enables to obtain the accumulated dispersion CD_k^S (and the slope CD'^S_k) of the span k inside the link (eq. 1). N is the number of spans in the link. The objective function (eq. 2) is the L1-norm distance with respect to CD_k^{typ} and CD'^{typ}_k, which are the (unknown) typical value of the accumulated dispersion (and slope) for each span k of the link. Their expression is given by eq. 3 where N_f is the number of possible fiber types. CD_j^{typ} and CD'^{typ}_j are the typical accumulated dispersion (and slope) of the fiber type j. $FT_{j,k}$ is a binary variable equal to 1 when the fiber type of span k is j (i.e., the fiber type index j varying from 1 to 7 as shown in Fig. 1). The searching space for CD_k^S and CD'^S_k is identical to the one used in step 1 (see Fig. 1).

The probability of the i[th] solution P_i (Table 1, eq. 4) is the product of two terms: one for the dispersion and one for the dispersion slope. Each one is obtained by multiplying the probability of all spans k of the link. The probability for each span is a gaussian function centered on the typical CD value (CD_k^{typ}) with a standard deviation σ_k given by the CD fiber specification. We have a similar expression for the CD slope. Each solution $\{CD_k, CD'_k\}_{k=1\ldots N}$ is minimizing the L1-norm (eq. 2), but the probability P_i is not necessarily optimal since the value of σ_k and σ'_k are not present in the objective function (not implementable in linear programming) and because we don't have a biunivocal relation between probability function and the objective function. The probability is then optimized afterwards once each solution is found and this optimal probability can be theoretically obtained. With the help of eq. 1, the derivatives of the probability P_i with respect to $CD_{k=1\ldots N}$ give a set of N equations with N unknowns (CD_k). The expression for the CD_i derivative is given by eq. 6 in Table 1. We have a similar expression for the dispersion slope probability. Solving this set of equations leads to CD_k and CD'_k values, maximizing the probability P_i. The normalized probability of each solution *i* with respect to the number of solutions N_{sol} is given by the expression of R_i (Table 1, eq. 5).

3. Simulation and live network results

We start by illustrating our method using a small network extracted from the production network consisting of 3 nodes, and 2 links according to the topology in Fig. 2. We have 6 lightpaths traveling over link 1, link 2 or link1+link2. The ground truth for the heterogenous link 1 is denoted as 37337 following the numbering of fiber types used in Fig. 1. Three cases are investigated (1, 2 and 3 in Fig. 1), where the characteristics of all fibers of this sub-network are the typical values (case 1); where they slightly deviate (case 2) or strongly deviate from these values (cases 3).

Table 2 reports the normalized probability R_i of the main solutions for the 3 cases. The maximal number of different fiber types is fixed to 2, the most probable practical case. For case 1, the most probable solution (R_i = 31%) corresponds to the true link composition. For case 2, the true solution is slightly less probable (at 26%) than the optimal found (at 27%). However, the 6 most probable solutions are permutations of the true solutions for cases 1 and 2. The inaccuracy of the CD measurements ($\Delta CD_j = 100$ ps/nm) prevents the algorithm to remove the permutation ambiguities. For case 3 (less likely than case 2 in deployed networks), the 2 SSMF are well identified but the 3 LEAF are replaced by 3 TWC. We propose to merge all permutations in a single solution whose probability is the sum of the probabilities of each permutation if we cannot rely on a very accurate CD measurement. The identification of the 3LEAF/2SSMF is 100% identified for the most realistic cases (1&2). For case 3, we have an ambiguity between 3/TWC/2SSMF (59%) and 3TWC/2SSMF (24%). This incorrect identification of fiber type was facilitated by the proximity of the CD and CD slope of the LEAF type (see the gray square in Fig. 1) and of TWC type. However, the computed CD/CD' values removed the TWRS type from the list of likely solutions.

Table 2 Solution accuracy for the 3 tested cases (green cells: true solution). The fiber type index and its color are shown in Fig. 1.

Sol	*Case 1*	*Case 2*	*Case 3*
1	**37337 (31%)**	73337 (27%)	74744 (20%)
2	73733 (29%)	**37337 (26%)**	47744 (17%)
3	73337 (17%)	33773 (23%)	47447 (15%)
4	33773 (13%)	73733 (19%)	77722 (11%)
5	37733 (9%)	33377 (3%)	**37337 (7%)**
⋮	⋮	⋮	⋮
12	44774 (0%)	44774 (0%)	17177 (<0.1%)

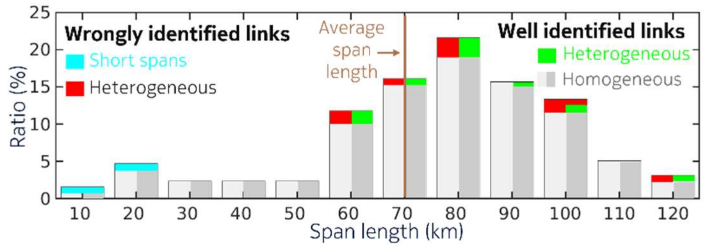

Fig. 3. Span length histogram. Prediction errors with the previous method (light gray, left) and generalized method (gray, right).

Table 3 Previous/generalized method comparison for the 2 links for which there is a mismatch between the previous method and client data. ∅ : no solution found.

Link Id	Configured Fiber Name	Homog. Analysis	Heterog. Analysis
20&38	4LEAF	∅	3LEAF / 1 TL
16&57	3LEAF/2SSMF	5TL	3LEAF /2SSMF

In this part, we consider CD measurements from the complete production network composed of 255 spans of fibers, 62 links, 148 lightpaths (from 8 to 903km). The span length histogram is represented in fig. 3 and highlights results of our generalized method compared to the previous one. Not surprisingly, the shortest spans remain ambiguous in terms of fiber types, but the new method is able to discriminate the fiber types over spans which were previously considered as not compliant, except in one link where we suspect mismatch with the client data. When compared to the operator's inventory, the total of correct predictions of the homogeneous analysis and the generalized method, are 87.1% and 97.7%, respectively. All homogeneous links are well identified with the previous and the generalized method. Links with different identifications between the previous and generalized method are shown in Table 3.

For the specific link id 20&38, the (heterogeneous) fiber type identification is different from the configured fiber name given by the network operator. A homogenous solution with 4 LEAF fibers with a chromatic dispersion of 5.8 ps/nm/km would be eventually possible but this value is much higher than the allowed range for LEAF type: no solution is returned by the previous method. The operator's inventory reports a (LEAF) homogeneous link which could also be attributed to either an under-estimation of the link length as high as 67km (i.e., 23% of the link length) or a quite large systematic error on all CD measurements (≈450 ps/nm). None of these scenarios are as probable as a mismatch between the field and the inventory. Our algorithm computed the TL type, which is unlikely in such a production network. Therefore, we can suspect it is a mix of LEAF and SSMF due to a splicing error. Knowing the exact position of the splice (by an Optical Time-Domain Reflectometer or OTDR) could help us to remove the potential ambiguity.

For the (heterogeneous) link id 16&57, a homogenous solution composed of 5 TL fibers with a dispersion of 9.37 ps/nm/km was found with the previous method. It corresponds to the highest allowed dispersion for TL type. The combination 3LEAF/2SSMF found by our generalized method is more probable and is confirming the client records.

4. Conclusions

By monitoring and correlating the accumulated chromatic dispersion of all network lightpaths, we identified the most probable fiber types in heterogeneous links with real data coming from a client production network. The solutions are ranked by likelihood. We obtained an excellent agreement (97.7% of success) with operator's inventory. For two links, our algorithm detected a potential error in the client network design.

This work has received funding from BPI France (grant 0168463/00) within the CELTIC-NEXT European project AI-NET Antillas. We thank P. Ramantanis, C. Delezoide, A. Legacy and S. Chenard for fruitful discussions.

5. References

[1] Y. Pointurier, "Design of low-margin optical networks," JOCN, Vol. **9**, no. 1, p. A9-A17 (2018).
[2] D.W. Boertjes *et al.*, "Practical considerations for near-zero margin network design and deployment," in JOCN, vol. 11, no. 9, 2019.
[3] T. Sasai *et al.*, "Digital Longitudinal Monitoring of Optical Fiber Communication Link," in JLT, vol. 40, no. 8, 2022.
[4] M. Eto *et al.*, "Fiber Type Identification based on Power Profile Estimation," ECOC, 2023.
[5] R. Shinzaki *et al.*, "Optical Network Tomography: Demonstration of Anomaly Loss Monitoring and Fiber-Type Identification", OFC, 2023.
[6] E. Seve *et al.*, "Fiber Type Identification: Alleviating Ambiguities," ECOC, 2022.