European Conference on Optical Communication (ECOC) 2022 © Optica Publishing Group 2022

Multilevel clustering in Point-to-Point Fiber Network Design

Simon Van den Eynde(1), Pieter Audenaert(1), Didier Colle(1), Mario Pickavet(1)

(1) IDLab, Ghent University - imec, Ghent, Belgium, pieter.audenaert@ugent.be

Abstract *We propose and test three metaheuristic approaches to extend a single-level FTTH network design heuristic to multiple levels. Each heuristic is evaluated on realistic graphs with over* 30000 *nodes and* 800 *terminals. We found a small but significant cost improvement. ©2022 The Authors*

Fig. 1: Atlanta dataset distribution level, connecting terminals to cabinets. Each color represents a cluster and cluster roots are indicated somewhat larger.

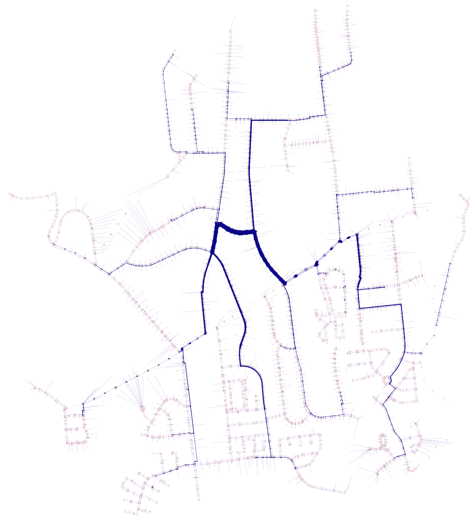

Fig. 2: Atlanta dataset feeder level, connecting cabinets to the central office. The more flow on an edge, the wider it is drawn.

Introduction

Bandwidth demand is ever-increasing and expected to keep increasing for at least 15 years [5]. One of the main ways to provide such bandwidth is through the installment of fiber to the home (FTTH). One of the major costs of a fiber network installment is labor, even more so in urban settings where fiber runs through trenched cables. These costs can be reduced by optimizing the network design planning.

A bottom-up fiber network design plan for a new location can be created as follows. Cluster the customers (terminals) in groups and assign a fiber cabinet to each group. Then, connect the fiber cabinets to a single Central Office, where connections are defined on a graph of nodes and edges based on the road architecture. While the number of levels could differ, this is a common approach to connect terminals, through a cabinet and distribution node to the Central Office. An example of a bottom-level connection can be seen at Fig. 1. For the same network the connection from the cabinets to the central office, see Fig. 2. Specifically, the problem tackled in this paper is the design of a multilevel FTTH network with capacities on some edges in a point-to-point fashion. That is, each terminal is connected with a fiber line directly to its cabinet; there are no splitters.

Finding the cost-optimal solution for a single cluster in a single level is a variant of a capacitated Steiner Tree problem which has been shown to be NP-complete and unapproximable in [2]. This paper focuses on large-scale instances with more than 5000 nodes, and therefore considers heuristic approaches. The FTTH network design problem is often solved approximately with MILP modeling such as in [4]. Other approaches have been used as well, such as genetic algorithms [6], [1] or through Binary Integer Programming [3].

However, all of these approaches consider smaller graphs and often require running times of 30 minutes and up. This makes it more difficult for telecom experts to interact with the system and to find good solutions on large graphs.

This paper proposes an iterative optimization metaheuristic that simultaneously optimizes root (cabinet/CO) locations and multiple levels of fiber routing. The metaheuristic provides a framework to extend a single-level FTTH network design heuristic to multiple levels, while still allowing for a fast evaluation on large graphs, requiring less than 15 minutes for a 32000 node, 800 terminal graph.

Background method

To optimize a single level, a custom iterative heuristic was implemented [8]. This is called the **base heuristic**. Its runtime complexity scales in practice about quadratically in problem size. However, the multi-level integration described in this paper can create improvements from any iterative heuristic.

The cost of a single level is quite complicated and has many variables, such as cable type, workers' hourly rates, etc. For this paper, and with the help of the software company Comsof, which specializes in FTTH network planning and design, these costs were summarized in three parts. 1) The fiber cost, i.e., total fiber length used for deployment multiplied by fiber per meter cost. 2) The trenching cost is the estimated cost for road works and is different from edge to edge. 3) The cluster cost, a base cost of 5000 to set up a cluster, e.g., a cabinet (this is an overestimation to artificially increase the number of terminals per cluster). The total cost of a level is then the sum of these three partial costs.

The calculation of the upper level depends on the level below. While the fiber cost has to be paid independently for each level, the trench cost only needs to be paid once per edge: when road works start for trenching on the bottom level, installments for the top level can happen simultaneously. Moreover, the division in clusters and subsequent cluster root locations become the terminals for the top level.

Method

Base Heuristic Iterate over different solutions for the bottom level by changing cluster compositions and updating fiber routing. Each iteration, based on the bottom layer cost, decides whether to keep the solution. After bottom layer optimization, optimize the upper layer.

Upper calculation (UC). The base heuristic can be improved by, in each iteration on the bottom level calculation, also running the base heuristic for the upper level. Then, the total cost can be used to decide if the new solution should be kept.

Adapted trenches at start (AT). Before the first iteration, the AT heuristic estimates the fiber locations of the upper level by running the base heuristic uncapacitated on all terminal nodes as if there is only one cluster (i.e., all terminals connect to a single cabinet). Then, based on the flow on each of the edges, they get a cost modifier. The higher the edge flow, the greater the cost reduction. Then, the base heuristic is run on the bottom level with the adapted edge costs, increasing the likelihood that it selects edges that are also good for the levels above. The inspiration for this approach comes from the perturbations used in [7] to optimize the Steiner Tree problem.

Adapted trenches at iteration (ATI). In each iteration of the base heuristic, if a new best solution is found (based on bottom level costs only), an upper-level solution is computed. Then, from this upper level, adapted trenches are calculated as described for AT. These adapted edge costs are then used for further bottom-level iterations.

The three methods require different amounts of coding, computation time, and quality of results. In the experiments, we will consider these three methods separately and the combination between UC and ATI, where the upper levels are calculated each iteration, and the edge costs in the bottom level are changed based on these adaptions.

Tab. 1: Dataset descriptions

	Atlanta	WashingtonDC
Nodes	6863	32109
Edges	10126	40224
Terminals	867	808

Experiments/simulation

The heuristic was tested on two realistic telecom fiber planning networks provided by Comsof. A short description is given in Tab. 1. For both datasets it is not hard to find a solution, only a limited number of edges has strong capacity restrictions, in line with common telecom deployments. Thus, the goal of the heuristic is not so much to find a solution but to find the most cost-effective solution. To facilitate comparisons and calculations, only two levels are built: a bottom level and an upper level. After a preliminary experiment series, it appeared that this setup resulted in a network design with a large number of small clusters. While this might be good theoretically, it is not realistic, so a material cost was added to each cluster of 5000. This resulted in more stable cluster numbers. For all experiments, 25 seeds of each map are generated, differing the demands on the terminals and the capacities on the edges, only

Tab. 2: Multilevel clustering costs for each algorithm for the Atlanta dataset. All costs are divided by 1000. The total cost is cost bottomlevel + cost upperlevel - overlapping trench cost.

Algorithm	bottomlevel cost	upperlevel cost	overlapping trench cost	total cost	clusters in bottomlevel	time (s)
UC + ATI	2205	7224	322	**9107**	33	149
ATI	2344	**7214**	**388**	9170	**29**	149
UC	**2081**	7298	227	9151	40	146
AT	2167	7314	239	9241	30	140
Base	2183	7325	214	9294	30	**118**

Tab. 3: Multilevel clustering costs for each algorithm for the WashingtonDC dataset. All costs are divided by 1000. The total cost is cost bottomlevel + cost upperlevel - overlapping trench cost.

Algorithm	bottomlevel cost	upperlevel cost	overlapping trench cost	total cost	clusters in bottomlevel	time (s)
UC + ATI	**5830**	54482	988	**59324**	210	908
ATI	6998	53929	990	59937	126	674
UC	5837	54725	**1010**	59553	211	830
AT	6769	**53822**	955	59637	**124**	219
Base	6610	53982	899	59693	**124**	**144**

the averages over all seeds will be shown. For each experiment, the bottom level and upper level are calculated.

Results and discussion

Tab. 2 and 3 present the results respectively on the Atlanta dataset and on the WashingtonDC dataset. For each experiment, the average costs of the bottom level, upper level, and total cost are displayed together with the number of clusters in the bottom level and the time in seconds. The overlapping trench cost comes from edges trenched in both the bottom and the upper level.

We expect that the base method has the best bottom-level cost because it only considers the cost at that level. However, in both tables, the base heuristic has not the best cost. This is due to the varying number of clusters at this level. Generally, by adding more clusters, the bottom-level cost decreases, and the upper-level cost increases. When running the bottom-level optimization with the base heuristic, we still want a realistic number of clusters, so this is artificially kept low (with the addition of cluster cost). What can be seen in both tables is that all evaluated methods have a higher overlapping trench cost, thus more overlapping trenches and fewer road works, compared to the base method.

For both datasets, the method UC + ATI has the best total cost and takes the longest to calculate. The UC method takes second place in total cost and time duration. It takes much more time to find a solution in the WashingtonDC dataset compared to the Atlanta dataset. This is due to the base heuristic scaling with the number of nodes as well as the number of terminals. And while the number of terminals is similar between both datasets, this is not the case for the number of terminals in the upper level, which is recalculated every bottom level iteration with UC.

Note that the additional time requirements are quite large compared to the gains in total cost. However, for such large problems, the extra computing time can still be a lot cheaper than the fiber installments.

Conclusions

In this paper, we propose three metaheuristic approaches to optimize the multi-level FTTH network design problem and compare them against a custom single-level heuristic. We find that the metaheuristic provides a small but significant gain in quality at the cost of running time. Since fiber installment can be quite expensive, the metaheuristic improvement can be a valuable extension.

The current version only considered two levels of fiber routing. While the metaheuristic should easily extend to multiple levels, we want to verify this and evaluate the effects on cost as well as time.

Acknowledgements

This research was partly funded by the Ghent University IOP project “Modelling Uncertainty in Hub Location Planning through Interdisciplinary Research” as well as by the imec/VLAIO project “Comsof Autonomous Planning Agent”.

References

[1] M. J. Barth, L. Mengue, J. V. Canto dos Santos, J. M. da Silva, M. J. Telles, and J. L. V. Barbosa. Multilevel Optimization Applied to Project of Access Networks for Implementation of Intelligent Cities. In V. Nazário Coelho, I. Machado Coelho, T. A.Oliveira, and L. S. Ochi, editors, *Smart and Digital Cities: From Computational Intelligence to Applied Social Sciences*, Urban Computing, pages 99–116. Springer International Publishing, Cham, 2019. ISBN 978-3-030-12255-3. doi:10.1007/978-3-030-12255-3_7.

[2] C. Bentz, M.-C. Costa, and A. Hertz. On the edge capacitated Steiner tree problem. *Discrete Optimization*, 38:100607, 2020. ISSN 1572-5286. doi:10.1016/j.disopt.2020.100607.

[3] L. Cen, K. Poon, A. Ouali, and M. Liu. Model Transformation for Automatic Design of GPON/FTTH Network. In *2021 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, pages 1382–1386, Dec. 2021. doi:10.1109/IEEM50564.2021.9672797.

[4] M. Chardy, M. C. Costa, A. Faye, and M. Trampont. Optimizing splitter and fiber location in a multilevel optical FTTH network. *European Journal of Operational Research*, 222(3):430–440, Nov. 2012. ISSN 0377-2217. doi:10.1016/j.ejor.2012.05.024.

[5] J. A. Hernandez, R. Sanchez, I. Martin, and D. Larrabeiti. Meeting the Traffic Requirements of Residential Users in the Next Decade with Current FTTH Standards: How Much? How Long? *IEEE Communications Magazine*, 57(6):120–125, June 2019. ISSN 1558-1896. doi:10.1109/MCOM.2018.1800173.

[6] R. Masoud, M. Ravanbakshsh, H. Farrokhi-Asl, and M. Taheri. Using metaheuristic algorithms for solving a hub location problem: Application in passive optical network planning. *International Journal of Supply and Operations Management*, 4(1):15–32, 2017.

[7] T. Pajor, E. Uchoa, and R. F. Werneck. A robust and scalable algorithm for the Steiner problem in graphs. *Mathematical Programming Computation*, 10(1):69–118, Mar. 2018. ISSN 1867-2949, 1867-2957. doi:10.1007/s12532-017-0123-4.

[8] S. Van den Eynde, P. Audenaert, D. Colle, and M. Pickavet. FTTH network design: An iterative Steiner tree heuristic. *[Submitted for publication to Journal of Heuristics]*, page 23, 2022.