On Location of POPs in FTTH Networks using Center of Gravity

Rasmus H. Nielsen, Tahir M. Riaz, Michael Jensen, Jens M. Pedersen and Ole B. Madsen
Center for Network Planning, Aalborg University, Denmark
email: {rhn, tahir, mj, jens, obm}@control.aau.dk

Abstract—This paper considers the cost of deploying a FTTH access network with varying number and location of supply points (POPs). A center of gravity optimization is used to compare the use of Euclidean distance and road distance along with a third combination of the two. Optimization is performed on the total line distance to the demand points (NTs). Initial locations are considered along with reduction in complexity. The results prove the combination method, that uses the optimized location from the Euclidean optimization as initial locations for the road optimization, to be the best. Substantial savings are achieved compared to locations in the existing copper network.

Index Terms—Communication Network Planning, ICT-infrastructure, Deployment Cost, Facility Location, Next Generation Network Technologies, Center of Gravity.

I. INTRODUCTION AND BACKGROUND

NEW ICT networks are currently being considered in many parts of the world, as copper lines are being replaced with fiber-optic lines even in the access networks - Fiber to the Home (FTTH). A large part of the involved network providers are new players in the field of data networks and have as such only little experience in this branch of network planning\(^1\). Nevertheless massive investments have already started the deployment of completely new fiber-based networks, expected to be the base of a new ICT infrastructure for the next 30 years or more. This motivates the research in planning of networks and, in particular, it makes it highly relevant to consider the methods used.

In planning of green field ICT infrastructures, three main elements are roughly to be considered; network terminations (NTs), points of presence (POPs) and fiber lines. The NTs - make up the fixed majority of the points in the network and the POPs make up a much smaller part of the points, which provides the services. These two sets of termination points are connected through the new ICT infrastructure consisting of fiber lines. As the set of NTs is known beforehand or a subset can be chosen with respect to marketing or technical preferences, the number of POPs and their locations are left to be considered. In case of existing network providers, new POPs are likely to be co-hosted with existing equipment and even in case of new network providers some preferences as to location are likely to be present. However, most new network providers are interested in optimizing on the number and location of POPs.

The locations of the POPs are of interest when minimizing the length of the fiber lines needed to connect the NTs. The better the POP locations in relation to the NTs, the less fiber lines will be needed. However, the POPs come at a cost as well and the number of POPs used is thus also of high importance.

This paper has been subject to only limited research in the context of network planning. Ref. [2] considers locations for active nodes in existing copper networks and [3] gives a thorough overview of the economics of multiple fiber-related technologies, but does only consider a simple representation of NTs and possible traces without real geographical data. Research on the subject is much needed as e.g. former research [4] has found it necessary to use existing POPs in the copper network for a future FTTH network.

This paper will contribute to the field of network planning by considering the number and location of POPs compared to the length of the lines needed to reach the NTs. Focus will be on FTTH access networks and distances, but the principle is transferable to e.g. higher network levels and traffic considerations. A center of gravity approach will be considered and evaluated with respect to the use of Euclidean and road distance for optimization. Further, the start conditions are considered along with a combination of the two optimization approaches as to reduce the computational complexity. A case study is used to test and compare the different approaches, from which the results are presented and compared. Before the conclusion, some aspects concerning optimization of POP locations are discussed along with further research.

II. PRELIMINARIES

In general the problem of finding locations for supply points in relation to a set of demand points is well-considered and is known as the facility location problem [5] [6]. With notation, \( D = \{d_1, ..., d_m\} \) as the set of \( m \) demand points supplied by the set \( S = \{s_1, ..., s_n\} \) of \( n \) supply points, the following will consider the NTs the demand points, \( D \), and the POPs, the supply points, \( S \). A distance function \( \delta(d_i,S) \) giving the shortest distance from \( d_i \) to the closest element in \( S \), is defined for optimization purposes.

In most cases the facility location problem is NP-hard [7], especially when considering the number of supply points as a variable. If considering every demand point as a possible supply point, the number of combinations to test, in order to find the optimal num-

---

\(^1\) In Denmark the power companies have used around 90 million Euros in order to put down more than 11,000 km of fibers and tubes. The same companies are planning to invest an additional 1.3 billion Euros in ICT infrastructure including 74,000 km fibers and tubes to reach (homes passed) almost 1 million or 40\% of the Danish households [1].
The general location/allocation problem is relevant in various parts of network planning, but with respect to location of supply points only a few of the specific facility location subproblems are of interest. This mainly means the capacitated as well as the uncapacitated facility location problem (UFLP). This paper will focus on using approximated methods to obtain solutions. Still, a smaller data set of possible supply points is introduced, as the set \( R = \{ r_1, ..., r_o \} \) of road intersections with a degree of 3 or higher.

The notation will be similar to the notation used for the general facility location problem introduced in section II.

### A. Cost of Deploying

For cost comparison some practical prerequisites are made. The cost of equipment and the traces needed to be dug up to deploy the lines are considered independent of the number of supply points. Hence these two parameters can be disregarded for comparison and doing so leaves back two parameters to consider, namely, the price of a supply point and the price of the lines connecting each demand point to its nearest supply point. For the case study the cost is greatly simplified by considering a supply point to be equally priced regardless of the number of demand points serviced and a line to be equally priced regardless of the number of fibers in the used cable.

For the latter comparison the lines have been priced as a cost of 100 Euro/km and a supply point as a cost of 100,000 Euro. These prices cover the capital expenditure as well as the accumulated operating expenses over 30 year converted into its current value.

### B. Algorithm for Location of Supply Points

Minimizing the total distance, \( \Delta_{\text{sum}} \), can be seen as minimizing the total distance at each supply point in \( S \) and reformulating equation 1 gives:

\[
\Delta_{\text{sum}} = \min_{i=1}^{m} \sum_{i=1}^{m} \delta(d_i, S),
\]

where \( \delta(d_i, S) \) contains the \( p \) demand points nearest to supply point \( s_i \). Hereby, the global optimization problem can be seen as a number of more local optimizations which is steady for a fixed number of supply points.

Algorithm 1 is used as a simple local search optimization loop. It starts with the initial locations \( S_0 \) and loops until \( \Delta_{\text{sum}} \) has not changed compared to its previous value.

#### Algorithm 1 OptimizeLocation.

1. \( \Delta_{\text{sum,old}} \leftarrow \infty, \Delta_{\text{sum}} \leftarrow 0, S \leftarrow S_0 \).
2. while \( \Delta_{\text{sum,old}} \neq \Delta_{\text{sum}} \) do
   3. for all \( s_i \) in \( S \) do
      4. \(<\text{OptimizeLocation}(s_i)>)\)
   5. end for
   6. \( \Delta_{\text{sum,old}} \leftarrow \Delta_{\text{sum}} \)
   7. \( \Delta_{\text{sum}} \leftarrow \sum_{i=1}^{m} \delta(d_i, S) \)
3. end while

The actual optimization is taking place in \(<\text{OptimizeLocation}(s_i)>)\), which is a center of gravity algorithm. Even as this in general does not give optimal results, it has been found efficient, especially when considering the large data sets used and the low complexity compared to other possible optimizations.

Two different measures are considered in the following: the Euclidean distance and the road distance.

#### B.1 Euclidean Distance

When using Euclidean distances as a measure of distance between a demand point and its closest supply point, the traditional center of gravity approach is valid. In this case each optimization considers the new location’s coordinates \( \Omega_i \) to be equal to:

\[
\Omega_i(x, y) = \frac{\sum_{j=1}^{p} d_{i,j} (x, y)}{p}
\]

and all that is left after algorithm 1 is to relate the last coordinates of \( \Omega_i(x, y) \) to the closest element in \( R \).

#### B.2 Road Distance

When speaking of distance in network planning, the matter becomes slightly more complicated compared to using the Euclidean distance. Planning of traces in ICT access networks mostly uses the road network as base [9], why the previous approach is not likely to give the good optimizations in general.

Using the road network for distance measures makes it necessary to define the center of a road structure, as the element in \( R_i \) (the road structure closest to \( s_i \)) from which the total distance, \( \Delta_i \), to any demand point is smallest possible. From this the center of the considered road structure can be calculated using the brute-force approach presented in Algorithm 2.
Algorithm 2 RoadCenter.

1: \( \Delta_i = \sum_{j=1}^{p} \delta(d_{i,j}, s_i) \)
2: if \( \Delta_i \neq \Delta_i^\text{old} \) then
3: \( \Delta_i^\text{old} \leftarrow \Delta_i \)
4: for all \( r_i,k \) in \( \hat{r}_i \) do
5: if \( \Delta_i > \sum_{j=1}^{p} \delta(d_{i,j}, r_{i,k}) \) then
6: \( \Delta_i \leftarrow \sum_{j=1}^{p} \delta(d_{i,j}, r_{i,k}) \)
7: \( s_i \leftarrow r_{i,k} \)
8: end if
9: end for
10: end if

C. Initial Locations

Preliminary results showed that the fitness of the optimization depended greatly on the initial locations, \( S_0 \). To provide such good initial locations an overlay to algorithm 1 is added, which starts out with a very high number of supply points and then recursively removes points until the actual number of supply points is reached. By removing the supply points with the lowest number of demand points, the supply points closest to the demand points are kept.

D. Reduction of Complexity

Of the two distance measures used in subsection B, the road distance will be more computational heavy to use, as distances must be calculated using shortest paths algorithms such as Dijkstra’s. This is especially a problem when testing out many scenarios (multiple number of supply points or initial locations) or considering large areas in which case the Euclidean distance is much faster to calculate. A third approach is thus included, which takes the optimized locations from the Euclidean distance and use those as initial locations for the road distance optimization.

E. Case Study

The case study area has been chosen to the typical Danish rural municipality Hals, which is located in the northern part of Denmark and has been subject to multiple analyses in the field of network planning. For the simulation all households (8638) have been considered demand points and the road network has been used for planning line ducts. All data are available through geographic information system (GIS), which has been used for the computational processing.

Table 1. Number of supply points, \( n_{sp} \), used for the case study, with interval \( i \), number of road intersections, \( n_{ri} \), and demand points, \( n_{dp} \), per \( n_{sp} \)

<table>
<thead>
<tr>
<th>( n_{sp} )</th>
<th>( i )</th>
<th>( n_{ri}/n_{sp} )</th>
<th>( n_{dp}/n_{sp} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 10</td>
<td>1</td>
<td>1611 - 161</td>
<td>8638 - 864</td>
</tr>
<tr>
<td>10 - 20</td>
<td>2</td>
<td>161 - 81</td>
<td>864 - 432</td>
</tr>
<tr>
<td>20 - 50</td>
<td>5</td>
<td>81 - 32</td>
<td>432 - 173</td>
</tr>
<tr>
<td>50 - 100</td>
<td>10</td>
<td>32 - 16</td>
<td>173 - 86</td>
</tr>
<tr>
<td>100 - 1611</td>
<td>100</td>
<td>16 - 1</td>
<td>86 - 5</td>
</tr>
</tbody>
</table>

For cost analysis and initial locations, a combination of 40 different numbers of supply points, as seen from table 1, including other relative information, has been used for the simulation.

IV. Results

The results have been achieved using Euclidean and road distance along with the combination proposed for optimization of \( \Delta_{sum} \). The relevant numbers of supply points have been tested in descending order and the optimized locations have been used as initial locations, \( S_0 \), for the next iteration with a sufficient number of points removed. Doing so yields the values of \( \Delta_{sum} \) as seen from Fig. 1 for 1-100 supply points in a loglog plot. Furthermore the maximum distance has be included as a practical comparison measure as seen from Fig. 2.

For comparison the two current supply point configurations in the copper network have been included. Looking at Fig. 3, showing the cost (see subsection III) of deploying a network with a given number of supply points, the most cost efficient combinations can be found. The global minimum is at 7 supply points, at a cost of 1.8 million Euro, with 6-10 supply points below 2 million Euro. Of the three methods considered, the combination method gives the overall best results and is better than the road method for more than 10 supply points. Using the Euclidean distance gives the worst results, even though this method performs better.
than using the road distance for some smaller intervals with a high cost. Compared to the two copper network configurations, substantial savings are achieved.

V. Discussion

The algorithms used are capable of finding much more cost efficient locations for supply points compared to the current locations in the copper network, however, some reasons can be presented not to choose the most cost efficient number and location of supply points. Such reasons could be already existing housing facilities or a deployment plan, where a high number of supply points is desirable in order to be able to reach many areas fast with only little ducting. Such parameters should thus be considered and already known locations should be used as fixed values for the algorithm.

An aspect for further research is the possible introduction of redundancy to all or a selection of the demand points, as network breakdowns due to cable cuts are likely to occur [10]. This will increase the complexity as the supply points should be located according to two sets of demand points - each supply point will be the primary point to one set of demand points and a redundant point to another set. Optimizing to one set can cause an increase in distances for the other set.

Further research can consider the use of other heuristic approaches suitable for this kind of optimization problems, such as genetic algorithms, convex optimization or space filling curves. Altering and using such algorithms should be considered to give more accurate and consistent result using fewer computations. Using different norms for distance should also be considered for further research.

VI. Conclusion

Two methods for optimizing locations for supply points (POPs) in relation to demand points (NTs) were introduced, based on a center of gravity approach, using Euclidean and road distances respectively. A third method was proposed by combining the two former methods. A way to find better initial locations was introduced. The methods give approximated optimization, but it seems intuitively likely, that they will converge to a result close the global minimum.

The methods were tested on a case study area and the results showed the combined method to give the lowest cost followed by the road and Euclidean distance. Any of the methods gave significantly better results compared to what could be achieved by using existing locations in the current copper network.

References