ABSTRACT Realistic reference networks are critical for all techno-economic evaluations in the design of an optical network architecture. As this information was not publicly available for Ireland, we generated such a reference network. In this paper we elaborate on the challenges faced during the generation and present some statistics. We also show how this is used in the design of an optical architecture for the same country. The level of detail of the reference network needed varies according to the layer of the architecture that is being designed.

1. INTRODUCTION

Continuous growth in the amount of the data transfer within national and global networks over recent decades demands new infrastructures and data transfer technologies. In line with this, the goal of the DISCUS project [5] was to develop an end-to-end design for a network that can provide a high-speed broadband capability of at least three orders-of-magnitude greater than today’s network to all users, while being ultra energy efficient and environmentally sustainable, and remain economically viable.

Typically, the network of a country has three layers: access, metro, and core. A signal originating at an end-user (or source) traverses through the access layer, then the metro layer and finally the core layer before descending back to the target in the exact reverse order. Using the DISCUS architecture a Long Reach Passive Optical Network [2] is deployed in the access part, the metro part is eliminated and a Transparent Optical Core Network is deployed in the core part [5].

A reference network plays a significant role in techno-economic evaluations in the design of an optical network architecture (see for instance [1]). This information was not publicly available for Ireland, so we generated such a reference network. In this paper we describe the challenges faced during the generation and present some analytics (Section 2). We then show how this is used in the design of an optical architecture. The level of detail of the reference network needed varies according to the layer of the architecture that is being designed. In this paper we show how the original reference network is reduced while maintaining important properties for the design of the architecture. We also elicit geotypes from the structure of the reference network and discuss the importance of this computation in the design of the optical distribution network.

2. A MICROSCOPIC NETWORK FOR IRELAND

We combine OpenStreetMap (OSM) data with additional sources of information to come up with a network that contains nation-wide microscopic view down to the level of individual building (customer). The road network data comes from OSM. The customer data, however, is coming from GeoDirectory Ireland ¹, and the locations of Exchange Sites data has been provided by Eircom ².

The customer data (building data) is basically a list of latitude and longitude pairs. There are 2,189,121 customers from GeoDirectory. In Figure 1 we show the exchange site data from Eircom. As it can be observed, exchange sites are more sparse in rural areas. We remark that to handle the massive amount of different data we had to develop special techniques for storage and access.

As mentioned before, part of the challenge is to integrate the different data sources that we have. There are two steps in this process: (a) the association of location of buildings with nodes in the OSM road network (see Figure 2(a)), and (b) the association of customer to local exchanges. For the first association, we found in the road network the closest node to the building (see Figure 2(b)). In Figure 2(a), the blue lines are denoting the distance from the location of the building to the corresponding node. Once the customers and exchange-sites are located in the OSM map, we associate customers with their closest exchange-sites. If we use a simple approach where we iterate over all exchange-sites for computing the closest exchange-site then this task could computationally very expensive. Therefore, we decompose this problem into smaller subproblems to avoid traversing the entire

¹https://www.geodirectory.ie
²http://www.eircomwholesale.ie/Our_Network/
Figure 1. LE sites (a) in Ireland and (b) in Cork

Figure 2. Reference network: (a) Buildings mapped to the network (b) An LE with associated buildings

We carry out this decomposition by creating a small network for each exchange-site as follows:

1) Associate each customer with their closest exchange site taking into account the Euclidean distance and find the farthest customer per site.
2) Use the distance to the farthest customer as radius to create a bounding box with latitude and longitude coordinates per site.
3) For each exchange site extract the customer and road node data into separate text files within the bounding box.
4) Compute the actual road distances using Dijkstra’s shortest-path-tree algorithm [3] for each exchange site.

Doing these steps allowed us to divide the problem into smaller problems for parallel and distributed processing of the data. One issue with this approach was that several customers ended associated with more than one local exchange. In those cases, the association with the shortest road distance was chosen.

3. Analysis and Reduction of the Reference Network

We focus our attention now on the analysis of the reference network. In Figure 3(a), for each local exchange, we show the distance to the furthest customer. A point (x, y) in this curve means that the probability of having a value greater than x is y. As we observe, more than 90% of the local exchanges have their furthest customers within 10 kms. We recall that the locations of the buildings/customers does not exactly correspond to the locations of the road nodes with which they are associated in the reference network as the former coordinates came from a different data set. In Figure 3(b) we show, for each customer location, the distance to their closest road node. As it can be observed, more than 90% of the locations are within 100 meters of the closest road node.

The level of detail of the original reference network is appropriate for the optimisation tasks that needed to be carried out at the level of the access network, but it is too detailed for the optimisation tasks that needed to be carried out at other levels. For instance, the locations of the customers were irrelevant for the optimisation at the higher levels of the architecture so the corresponding nodes could be safely removed if they were not disturbing the connectivity of the network.

Our challenge was then to come up with a less detailed reference network without compromising the properties we were interested in keeping. On the one hand we wanted to ensure that distances between nodes were kept...
Figure 3. (a) Distances from local exchanges to their furthest customers (in kilometres) (b) Distance from a location to the closest road nod (in meters). A point (x,y) in these curves means that the probability of having a value greater than x is y.

while on the other hand our purpose was to ensure that any disjointness claim made on two paths of the reduced reference network was sound (i.e., to avoid claiming that two paths were disjoint when they were not).

Let us keep in mind that the reference network is an undirected graph. The first step in the reduction was to remove nodes of one degree that were not location of local exchanges. Notice that the elimination of these nodes and their corresponding links (as shown at the left hand side of Figure 4(a)) affect neither of the properties we were interested in keeping.

The second step was to remove nodes of degree 2 by replacing their incident links with a link whose length was equal to the sum of the removed links (see right hand side of Figure 4(a)). It is easy to show that this transformation does not disturb the properties that we want to enforce. As we are replacing the two links with a link of equivalent distance any path relying on the deleted links remains of the same length. Also, as there are no paths being merged by this reduction, the disjointness claims on the reduced reference network remain sound. The elimination of nodes of degrees 1 and 2 led to a significant reduction in the number of nodes and links as shown in Figure 4(a).

Removing nodes of degree 3 is not so straightforward. In Figure 4(b) we show one attempt to do so. Here node c and its incidents links are removed. For each pair of node we add link whose length correspond to the sum of the corresponding links. However, such reduction would lead to invalid disjointness claims. Notice that in the reduced graph it is possible to have two disjoint paths: one connecting the pair (a, d) and another connecting (b, d). However, in the original graph it is not possible to have such paths, which means that the reduction is not sound when it comes to disjointness.

Figure 4. Reduction rules on reference network

4. GEOTYPE CLASSIFICATION

In this section we use the reference network to classify exchange sites and customers into three geotypes: rural, sub-urban, and urban. Unlike previous work (e.g., [4]) that estimates the geo-types based on the number
The geo-type refers to the customer density in a given area. In this paper, we use the number of households per square kilometre to define the following three geo-types: Rural geo-type with up to 10 households in 1 km$^2$, Sub-urban geo-type with [11, 500] households in 1 km$^2$, and Urban geo-type with more than 501 households in 1 km$^2$. Figure 5(a) shows the geo-type distribution in Ireland, visually it can be observed that the majority of the country is rural or sub-urban areas. In total we calculated 26,253 Rural (black), 28,500 Sub-urban (grey), and 873 Urban areas (red).

In Figure 5(b) we focus our attention in areas with up to 100 customers in 1 km$^2$. Rural and Sub-urban geo-types dominate in Ireland covering 51% of the Irish population (Rural: 6% and Sub-urban: 45%), and urban areas cover 48% of the Irish population (i.e., about 1,054,177 customers). Additionally, we would like to remark that only 16 urban areas cover more than 4000 customers in 1 km$^2$. These areas are geographically distributed in Dublin, Cork, and Waterford. This information is consistent with the Irish central statistics office\(^3\), which indicates that the most populated cities are Dublin, Cork, Limerick, Galway, and Waterford, being Dublin and suburbs the largest urban areas in the country. Finally, we found that in Ireland 26 exchange sites were rural, 948 sub-urban, and 230 urban.

5. CONCLUSIONS
We have elaborated on the process followed to generate an Irish reference network. We discussed the approach followed to reduce the level of detail while ensuring that important properties of the network are kept. We also presented our approach to elicit geo-types from the structure of the reference network.

ACKNOWLEDGMENTS
This work was supported by DISCUS (FP7 Grant Agreement 318137), and Science Foundation Ireland (SF) Grant No. 10/CE/I1853. The Insight Centre for Data Analytics is also supported by SFI under Grant Number SFI/12/RC/2289.

REFERENCES

\(^3\)http://www.cso.ie