Are Optical Networks Scale-free?

Carol Meertens, Joost Pijnaker.

July 4, 2007

Abstract

Scale free graphs are graphs with a power law degree distribution. Many real-world networks like the WWW, email networks, metabolic networks or social networks have been shown to be scale-free [1, 2, 3, 4]. They also exhibit related properties like high clustering and being of 'small-world' type [5].

We question whether graphs of optical networks are scale-free. We assemble data of eight networks, map them to graphs and investigate their degree distribution. For one of the networks we consider different ways of modeling. Different models will result in different graphs and thus different degree distributions.

1 Introduction

Up until 1999 networks were simulated mostly by creating Erdös-Rényi random graphs. In Erdös-Rényi random graphs all possible pairs of vertices have the same, fixed probability of being inter-connected by an edge [6]. Because of this fixed probability of vertices being connected, the degree distribution of an Erdös-Rényi random graph is a Poisson distribution.

In 1999 and the years following that a lot of real-world networks were studied, along with their degree distributions [7]. It appeared that these real-world networks have degree distributions quite different from the Poisson distributions for Erdös-Rényi random graphs: there is a relatively low number of degrees close to the mean and a lot of degree values are either very low or very high. Figure 1 illustrates the difference between the degree distributions of Erdös-Rényi random graphs and those of real-world networks.

'Scale free graphs', by definition, are graphs having a power law degree distribution. Many real-world networks appear to have this distribution [7]:

$$P(k) \sim \frac{1}{k^{\gamma}} \tag{1}$$

Note that 'scale-free' is defined here quantitatively for *graphs*; not for the networks themselves[1]. When talking about scale-freeness of networks there often is a more or less implicit notion of mapping the network to a graph.

[1] Also note that there is no mention of *scalability* here. There is no immediate relation between scale-freeness and scalability.

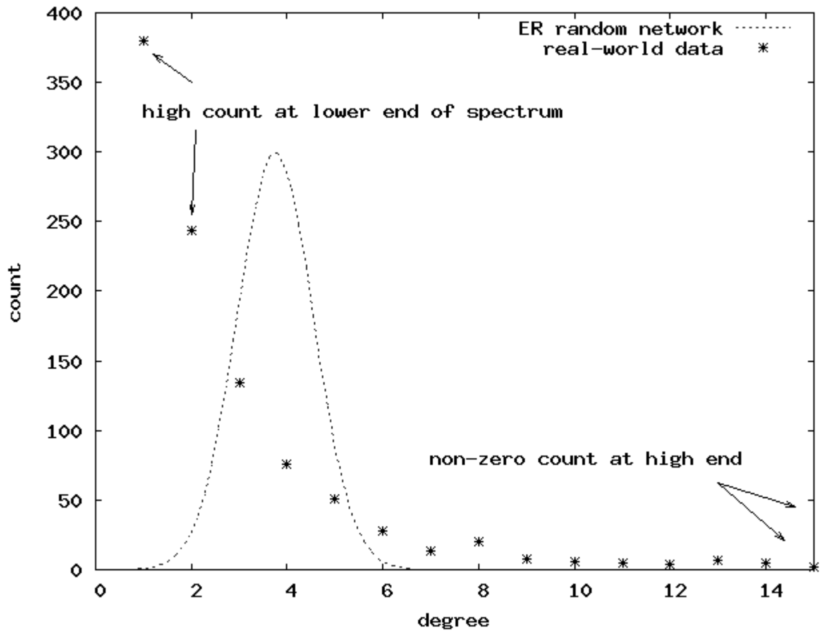


Figure 1: Real-world degree distributions vs. degree distributions of random graphs. The degree distribution of most real-world networks (dots) is not what one would expect from a random network (line).

This raises an important issue when scale-freeness for optical networks is examined. In this document an 'optical network' is defined as a multilayer network based upon optical, circuit switched, technology. An optical network is typically 'hybrid'. A network can often be mapped to a graph in multiple ways and in the particular case of optical networks the mapping to choose is not straightforward at all [8]. There are decisions to be made in the process: what will be the vertices? The physical devices? Their interfaces? And what will be the edges? The choice will certainly influence the topology of the graph. It is possible that one mapping will result in a scale-free graph while another mapping will not.

The Network Description Language (NDL) is a descriptive language for hybrid networks [9]. The applicability of NDL can be tested by generating networks in NDL format. These generated networks should have similar properties as real-world networks. Scale-freeness can be one of these properties.

Thus we come to the following key questions to answer:

- Are optical networks scale-free?
- Is the answer to the first question dependent on the way one models the network? And if so: how?

We assembled data of eight optical networks. Most data were supplied to us in the form of graphical maps. The data for the SURFnet6 network was supplied to us in NDL format. Using NDL we explored different models to create the graph. For each graph we decided whether it was scale-free by calculating the standard error in the scaling exponent γ from equation (1). We also generated graphs having either Erdös-Rényi random or scale-free degree distributions. The plots of these degree distributions were compared to those of the optical networks.

The rest of this document is organized as follows: in Section 2 we describe different models we used to graph the SURFnet6 network. Section 3 states our findings for both the graphical maps and the SURFnet6 data. We will end with Section 4, formulating the answers to the key questions stated earlier.

2 Models of Networks

Different models of a network result in different graphs and thus different degree distributions. It is possible some will be scale-free while others will not. Any computer network can always be mapped to a graph in multiple ways. But in the special case of *optical* networks, where connections typically cross layer and technology boundaries, the way one models the network is not necessarily straightforward.

We therefore describe four alternative models we used to map optical networks to graphs. The degree distributions that result from applying the models to our NDL data will be presented in Section 'Results'.

We used the NDL format as the basis for our models. Version 1 of NDL describes the network using four classes [10]. These classes are our candidates for mappings to vertices and/or edges:

Location A place where devices are located.

Device Any kind of machine that is connected to the network.

Interface The connection between the device and the rest of the network.

Link An (abstract) connection between two interfaces

We refer to Appendix A for an example network description in NDL format.

Model I: Locations

In this model we map Location elements from the network description of the optical network to vertices in the graph. There is an edge between the vertices whenever there are Interfaces within the Locations and their Devices being connected. In NDL this is done via the *connectedTo* property of an Interface (see Appendix A).

Two Locations can be linked multiple times through different Devices and/or Interfaces. In that case we can choose to either draw multiple edges or to draw only one edge between the two Locations. This leads to two different sub-models and thus two different degree distributions. When choosing to draw multiple edges in this case, the degree distribution will typically have more points and the degrees can extend to larger values (see Figure 2).

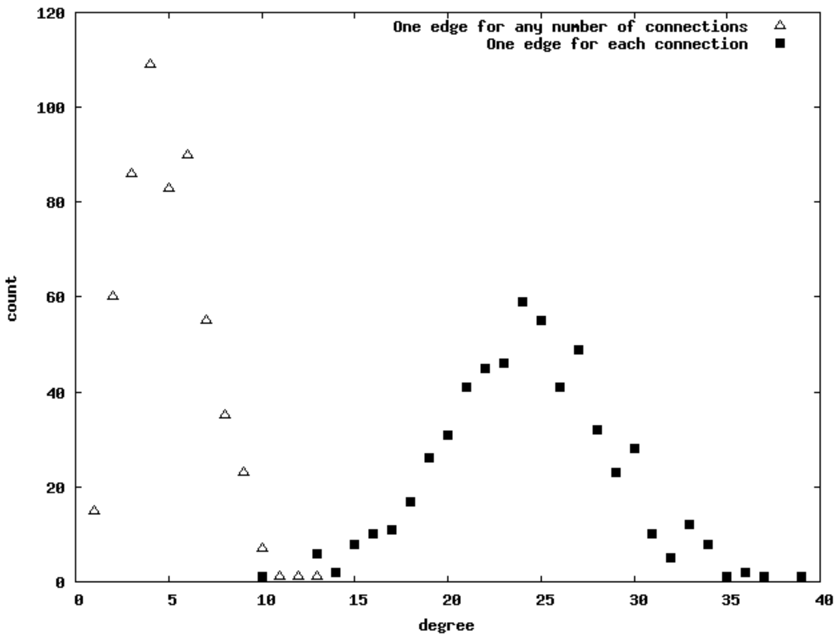


Figure 2: Example degree distribution for Model I in two versions: Location vertices either can (filled squares) or cannot (open triangles) be interlinked more than once, depending on whether multiple Interfaces are connected. When choosing to draw a separate edge for each connection, the distribution will typically have more points and degree values tend to be larger. These degree distributions are generated using the algorithm for Erdös-Rényi random graphs.

Model II: Devices

Here we map the Device class to vertices and connected Interfaces to edges. As with Locations pairs of Devices can be linked more than once via their Interfaces and the degrees of vertices. So the effect on the degree distribution will be the same as described in 'Model I: Locations'.

Model III: Interfaces

In this model we map the Interface class of the network description to vertices. An edge is drawn whenever Interfaces are connected. This will result in an ensemble of disconnected graphs, all consisting of a pair of two connected vertices. The degree distribution will show one point at degree value one. We therefore extended the model: two Interfaces are thought to be connected whenever they are part of the same Device. This way a Device holding 10 Interfaces can be represented by a piece of a graph having 10 vertices and 45 edges connecting the vertices full-mesh[2].

[2] Each of the 10 vertices is connected to nine others. Two connected vertices share an edge, so the total number of edges in the full-mesh is $\frac{1}{2} * 10 * 9 = 45$.

Another way to extend the model would be to add one vertex for each Device. Interfaces then are not thought to be linked to each other full mesh, but rather to the center of the Device they are part of. Figure 3 illustrates this. In this case the vertices that map from the Interfaces all will have a degree value of two. The vertices for the Devices will have the same degree as in model II. So the degree distribution of this sub-model will be the same as the degree distribution of model II, apart from a high degree two count. We conclude it would not be useful to apply this model to the NDL data separately.

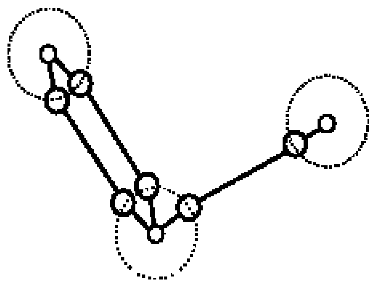


Figure 3: Vertices for Interfaces and for Devices. Dotted lines represent Device elements in the NDL data. Graph vertices are smaller circles.

Not all Interfaces within a Device need to be connected to an Interface within another Device. We can either choose to map them to vertices or to leave these out of the graph. Figure 4 shows what can happen to the degree distribution when incorporating the unused interfaces in the model. The degree of the *used* Interfaces will increase and the extra *unused* Interfaces will add to the degree counts. This could in theory have an effect on whether the distribution is a power law or not.

Model IV: Multilayer

The fourth and last model we consider does not take an NDL class for the vertices. Instead we make use of the *capacity* property NDL offers us [10]. The capacity property of the Interface class in our data corresponds to the number of optical channels that Interface supports. In this model we map both Interfaces and Devices to vertices. We map each SDH channel to an edge between a Device vertex and an Interface vertex: see Figure 5. We are deliberately crossing layer boundaries: edges in our graph explicitly do not belong to the same network layer here.

Interfaces in our NDL data have 48 or 192 channels. Degree values can therefore be very high and this will be visible in the degree distribution. Figure 6 shows what can happen to the degree distribution when mapping channels to edges within the Devices.

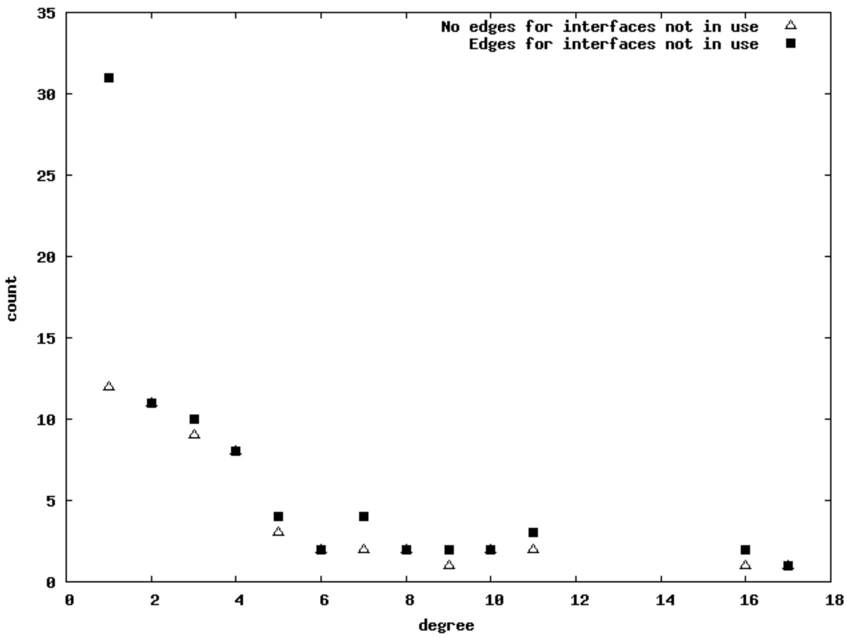


Figure 4: Degree distribution of Model III: with and without in-use Interfaces. The not-in-use Interfaces contribute to a high degree one count.

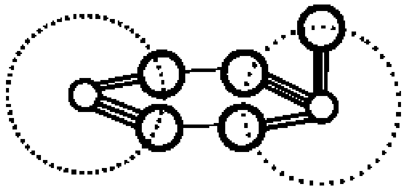


Figure 5: Model IV; edges can also be SDH channels.

Model V: Links

A Link is a connection between two Interfaces. As a consequence, when mapping the Link class to vertices, all vertices would have degree two. The degree distribution would thus have one point only: at degree two. Therefore we did not consider to apply this model to the NDL data given to us.

Generic Remark on the Models

Above we showed how different models can affect the graph and the degree distribution. Whether taking another model will change a distribution from power law to *not* power law (or the other way around) in most cases can only be determined by really parsing the concrete network description at hand.

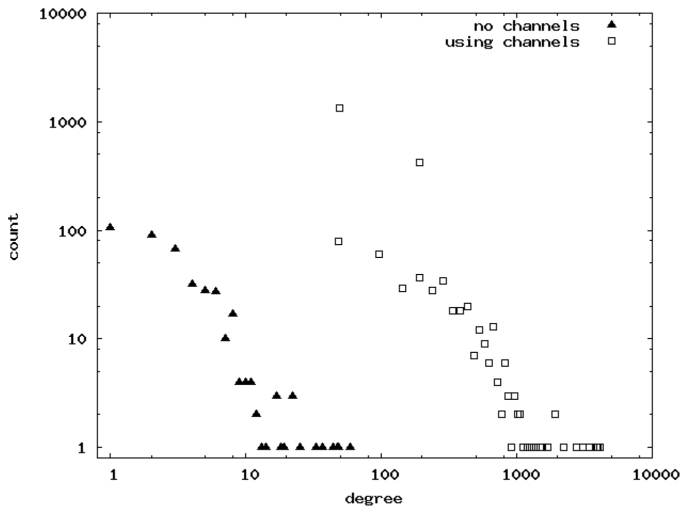


Figure 6: Model IV: the effect on the degree distribution when mapping channels to edges. To create this plot first a degree distribution was generated using the Barabási-Albert algorithm for scale-free graphs (closed triangles). Based on this first distribution a second distribution for Model IV was generated (open squares). Each edge from the first distribution was replaced by either 48 or 192 edges with probabilities of 75% and 25% respectively. Extra vertices with corresponding degrees were added for the Interfaces.

3 Results

We assembled data of eight optical networks. Seven of those we got in the form of graphical maps[3]. We analyzed these maps by mapping physical optical devices to vertices and fiber links to edges. The number of optical devices in the graphical maps was in the order of 100. The degree values of the vertices were counted by hand.

The remaining data set was an NDL file describing part of the SURFnet6 network[4]. There were 69 Device elements in the NDL file. We evaluated the data for three out of four models described above. We dismissed Model I as the data only contained one Location element.

For all the resulting degree distributions we calculated the scaling exponent γ along with its standard error. Table 1 shows the results. The standard error in γ was never below 5%, so the power law function is not a good approximation for any of the resulting degree distributions. As the power law function defines scale-freeness we cannot be conclusive about the scale-freeness of any the graphs.

Figure 7 shows one of the degree distributions for the SURFnet6 data. In this case Model II was used to map the network to a graph. The small number of points in this plot is related to the fact that optical devices in the SURFnet6

[3]See Appendix C for an example of the graphical maps.
[4]SURFnet6 is the Dutch research and education network.

data	γ			
SC02	0.9	$\pm$	0.2	(21%)
SC03	0.7	$\pm$	0.3	(44%)
SC04	0.58	$\pm$	0.14	(23%)
SC05	0.90	$\pm$	0.16	(18%)
SC06	1.12	$\pm$	0.11	(10%)
GLIF	0.86	$\pm$	0.17	(20%)
Internet2	2.2	$\pm$	0.6	(27%)
SURFnet6 model II	1.0	$\pm$	0.4	(40%)
SURFnet6 model III	0.0	$\pm$	0.4	(-)
SURFnet6 model IV	1.1	$\pm$	0.3	(26%)

Table 1: γ values for the degree distributions. See Section 'Data Sources' at the end of this document for pointers to the raw data files.

network have a fixed and rather small number of interfaces[5]. The degree values in the graphs likewise fall within a small range. This is the case for all networks and models we examined.

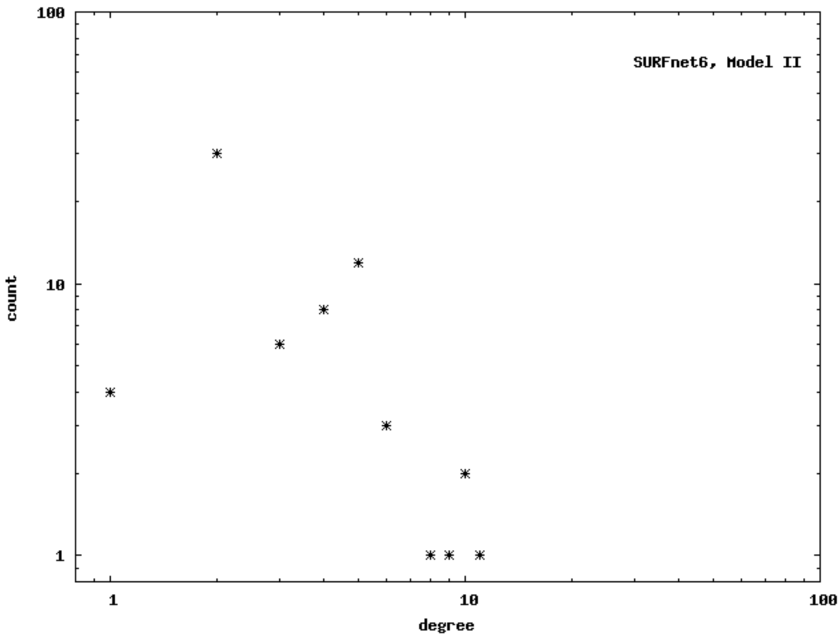


Figure 7: SURFnet6, Model II

Because our degree distributions can not be said to be approximated by the power law, we tried a qualitative comparison to both scale-free and random graphs. We generated nine degree distributions using the Erdös-Rényi algorithm (for generating Erdös-Rényi random graphs, [6]) and nine degree distributions using the Barabási-Albert algorithm (for generating scale-free graphs,

[5]An optical device typically has in the order of ten ports an not all will be in use. A new optical device generally will be installed before all ports are being used.

[11]). Both algorithms were parametrized such that the total number of vertices would match the number of Device elements in our NDL data.

The NDL data could not be visually matched with either the random set or the scale-free set. Therefore it was not possible to make a qualitative statement about the scale-freeness of the NDL data using Model II. Appendix C compares nine runs for both algorithms with the degree distribution for the NDL data.

4 Conclusion

To answer the question whether optical networks are scale-free we examined the degree distribution of eight optical networks. One of those we analyzed using three models that map the network to a graph.

The standard error in the scaling exponent γ in our case was never below 5%. We therefore could not be conclusive about the scale-freeness of optical networks. Using different models to create the graphs did not affect this. A qualitative comparison of one of the degree distributions to the degree distribution resulting from simulations also yielded inconclusive results.

Degrees in our graphs cover no more than one order of magnitude. This is related the fact that the optical devices in our networks have a limited number of ports.

References

[1] A. Barabási, R. Albert, and H. Jeong. Scale-free characteristics of random networks: the topology of the World Wide Web. *Physica A: Statistical Mechanics and its Applications*, 281:69–77, 2000.

[2] H. Ebel, L. Mielsch, and S. Bornholdt. Scale-free topology of e-mail networks. *Physical Review E*, 66:035103, 2002.

[3] H. Jeong, B. Tombor, R. Albert, Z. Oltvai, and A. Barabási. The large-scale organization of metabolic networks, 2000.

[4] X. Guardiola, R. Guimera, A. Arenas, A. Diaz-Guilera, D. Streib, and L. A. N. Amaral. Macro- and micro-structure of trust networks, 2002.

[5] M. E. J. Newman. Random graphs as models of networks. Working Papers 02-02-005, Santa Fe Institute, February 2002. available at http://ideas.repec.org/p/wop/safiwp/02-02-005.html.

[6] P. Erdös and A. Rényi. On random graphs. *Publicationes Mathematicae*, 6:290–297, 1959.

[7] A. Barabási and E. Bonabeau. Scale-free networks. *Scientific American*, pages 60–69, May 2003.

[8] F. Dijkstra, B. Andree, K. Koymans, J.v.d. Ham, and C. de Laat. A multi-layer network model based on ITU-T G.805. May 2007. to be published.

[9] Jeroen van der Ham, Freek Dijkstra, Franco Travostino, Hubertus Andree, and Cees de Laat. Using rdf to describe networks. *Future Generation Computer Systems, Feature topic iGrid 2005*, 2006.

[10] J.v.d. Ham, P. Grosso, R.v.d. Pol, and A. Toonk C. de Laat. Using the network description language in optical networks. May 2006.

[11] R. Albert and Barabási. Emergence of scaling in random networks -. *Science*, pages 509–512, 1999.

Data Sources

All retrieved 15th June 2007:

SC02 through SC06: `http://scinet.supercomp.org/`

GLIF: `http://www.glif.is/publications/maps/GLIF_8192-03August2005.jpg`

Internet2: `http://paintsquirrel.ucs.indiana.edu/fiber_map.pdf`

SURFnet6: not public

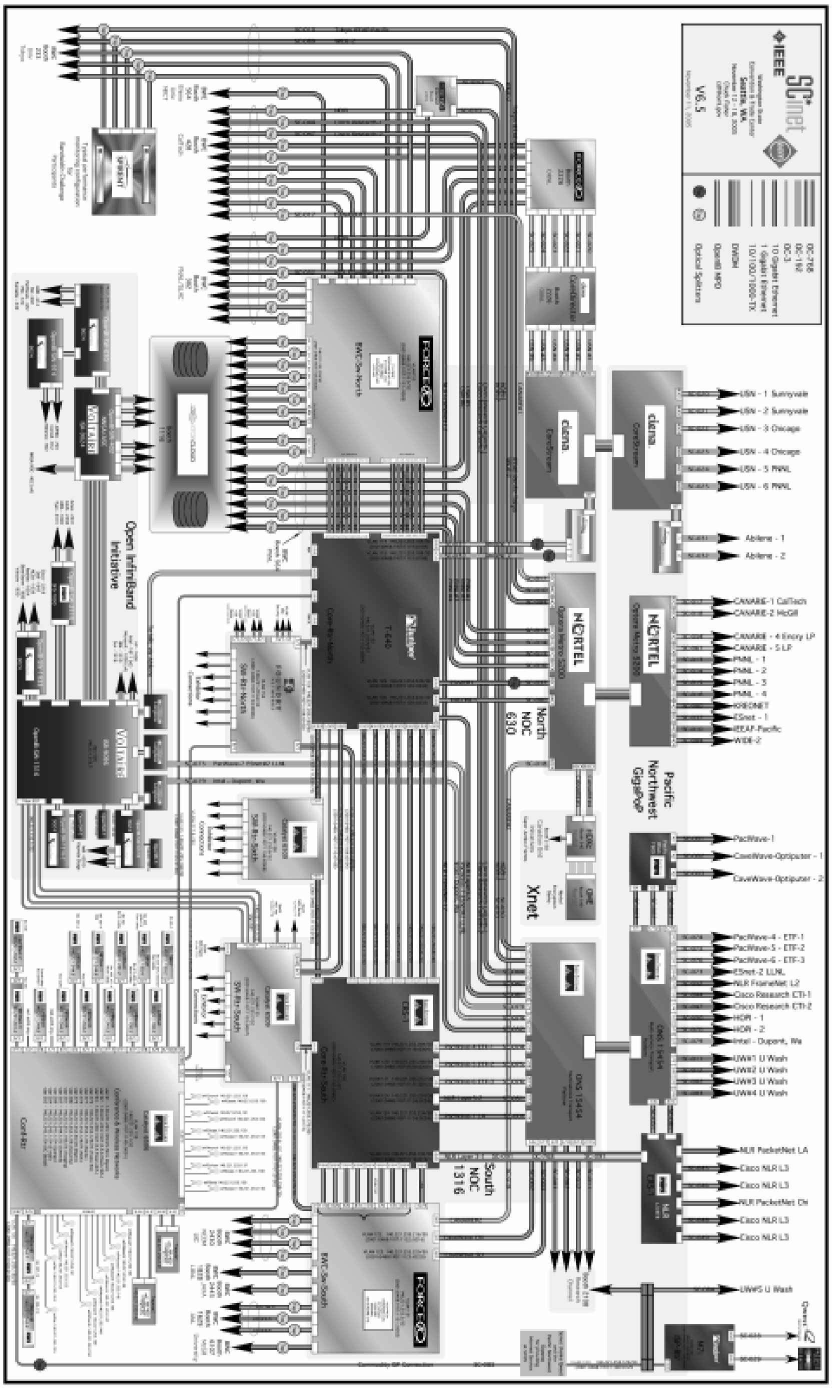
Appendix A: example NDL

This is an example network description using the Network Description Language (NDL, [9]). This is here for clarifying the way we modeled the SURFnet6 data; not in any way to show a complete or even correct piece of NDL. Please refer to `http://www.science.uva.nl/research/sne/ndl/` for up-to-date and complete information about NDL and for more examples.

```
<?xml version="1.0" encoding="UTF-8"?>
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
        xmlns:ndl="http://www.science.uva.nl/research/sne/ndl#">
<ndl:Location rdf:about="#Netherlight">
        <ndl:name>Netherlight Optical Exchange</ndl:name>
</ndl:Location>
<ndl:Device rdf:about="#tdm3.amsterdam1.netherlight.net">
        <ndl:name>tdm3.amsterdam1.netherlight.net</ndl:name>
        <ndl:locatedAt rdf:resource="#amsterdam1.netherlight.net"/>
        <ndl:hasInterface rdf:resource="#tdm3.net:501/1"/>
...
</ndl:Device>
<ndl:Interface rdf:about="#tdm3.amsterdam1.netherlight.net:501/1">
        <ndl:name>tdm3.net:POS501/1</ndl:name>
        <ndl:connectedTo rdf:resource="amsterdam1.net:5/1"/>
        <ndl:capacity>1.2E+9</ndl:capacity>
</ndl:Interface>
...
</rdf:RDF>
```

Appendix B: example map

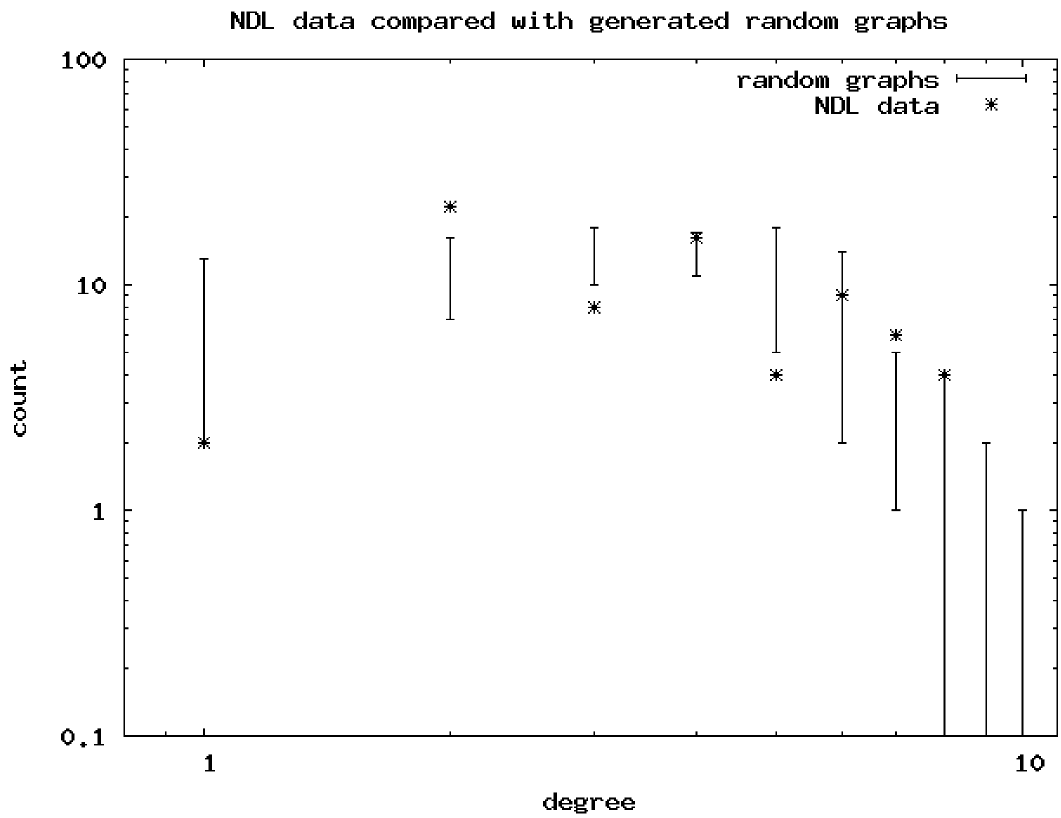
Most of the network information available to us was in the form of a graphical map like the one shown here. This map is retrieved 15th June 2007 from `http://scinet.supercomp.org/`.

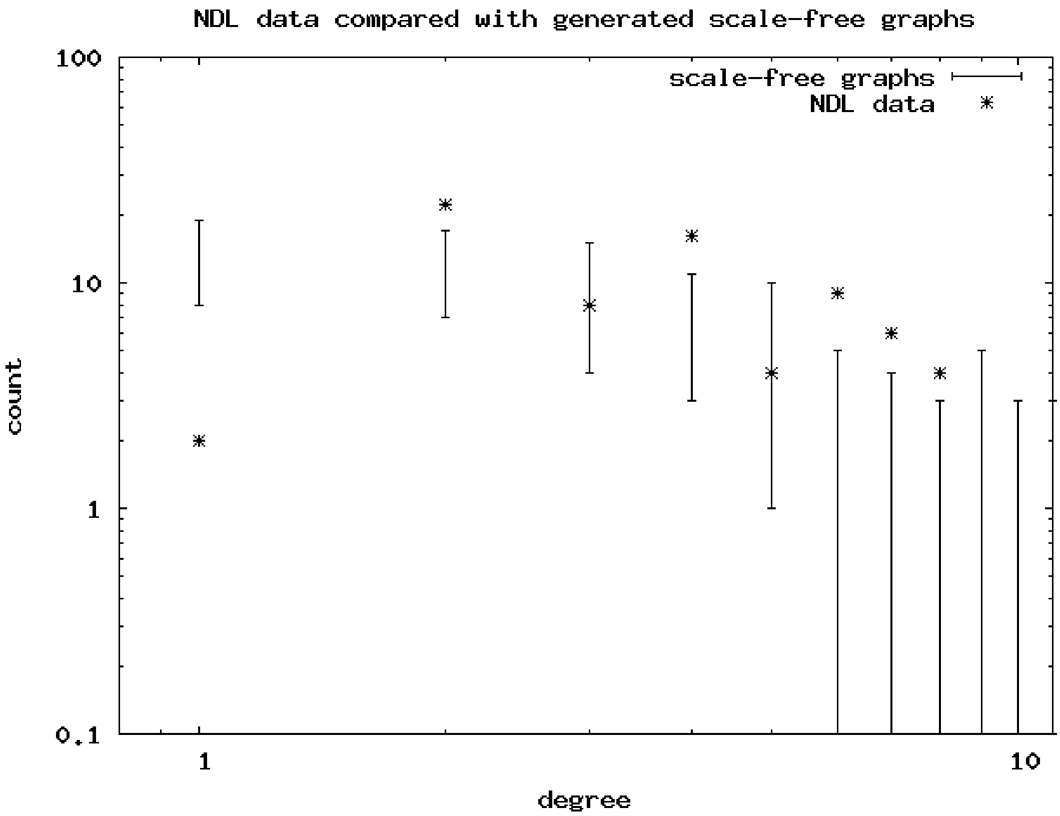


Appendix C: degree distributions from simulated networks

Below is a qualitative comparison of the NDL data to both scale-free and random graphs. We generated nine degree distributions using the Erdös-Rényi algorithm (for generating Erdös-Rényi random graphs, [6]) and nine degree distributions using the Barabási-Albert algorithm (for generating scale-free graphs, [11]).

Both algorithms were parametrized such that the total number of vertices would match the number of Device elements in our NDL data. Vertical bars extending to the bottom appear whenever there was a run having zero vertices with a certain degree.





NDL data compared with generated scale-free graphs
scale-free graphs
NDL data
count
degree
100
10
1
0.1
1
10