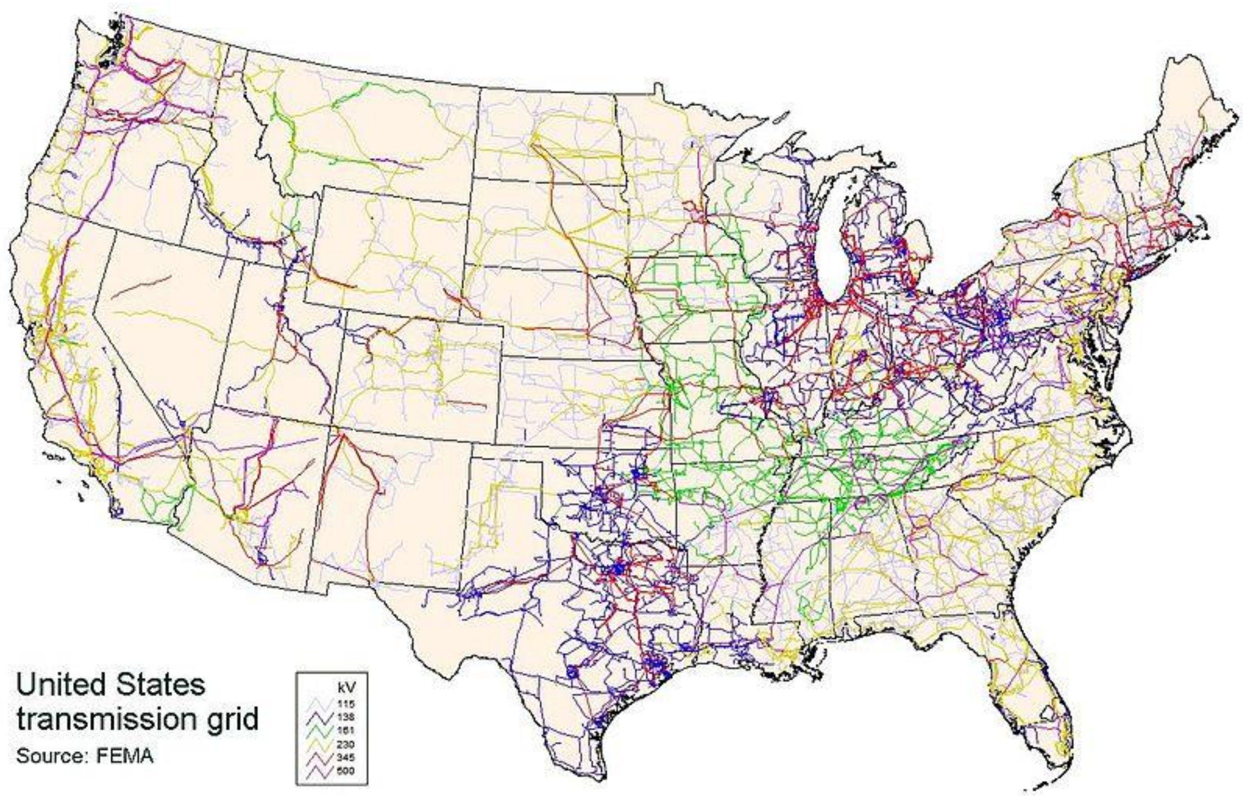
Analysis of network

Power Grid US

Diego Angulo Ramirez



Analysis of the network "Power Grid US"

This **undirected** network contains information about the power grid of the Western States of the United States of America. An edge represents a power supply line.

A node is either a generator, a transformator or a substation.

We are not certain about the exact year where the data was taken, but we are sure is sometime before to 1998, most probably 1996.

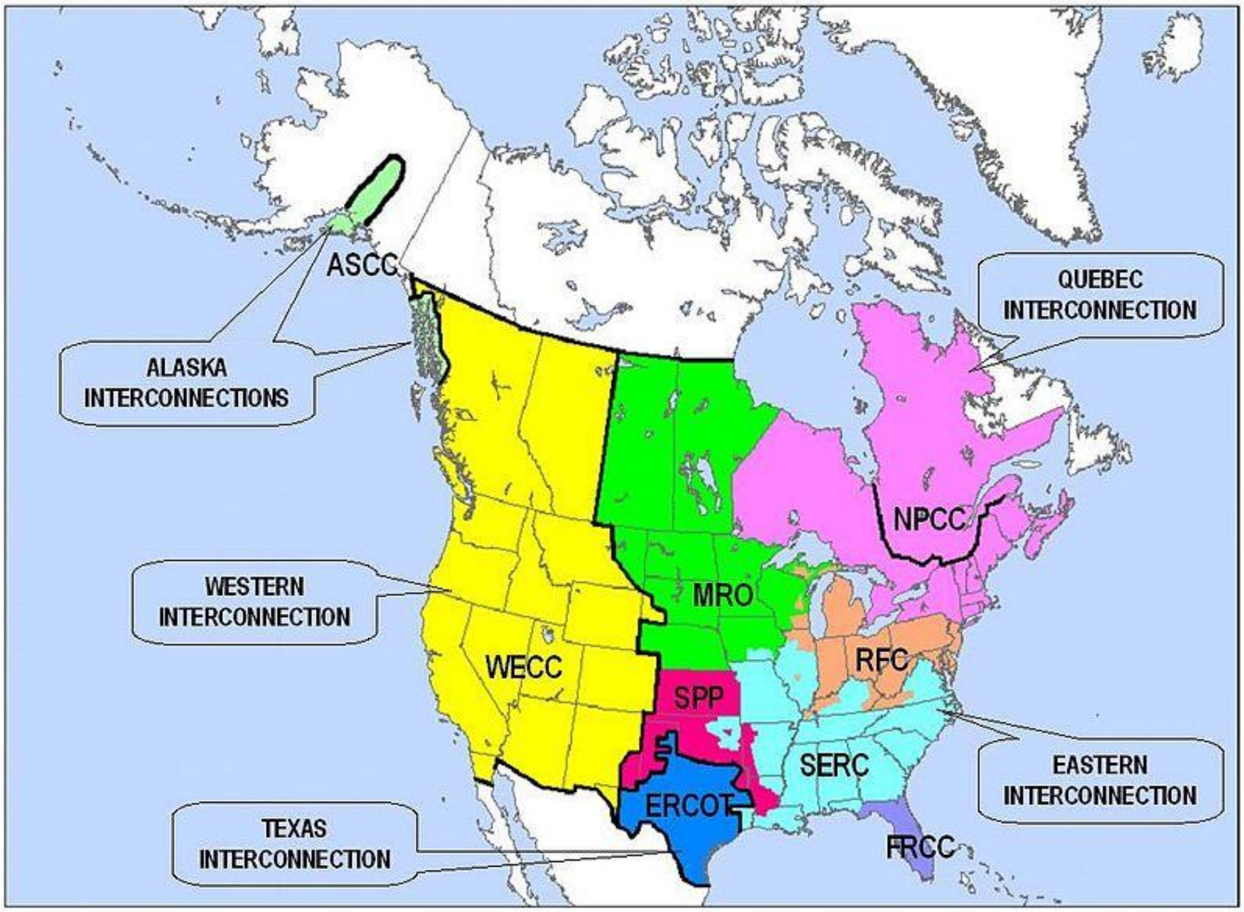
The dataset used was compiled by D. Watts and S. Strogatz

Collective dynamics of ‘small-world’ networks, D. J. Watts and S. H. Strogatz, Nature 393, 440-442 (1998).

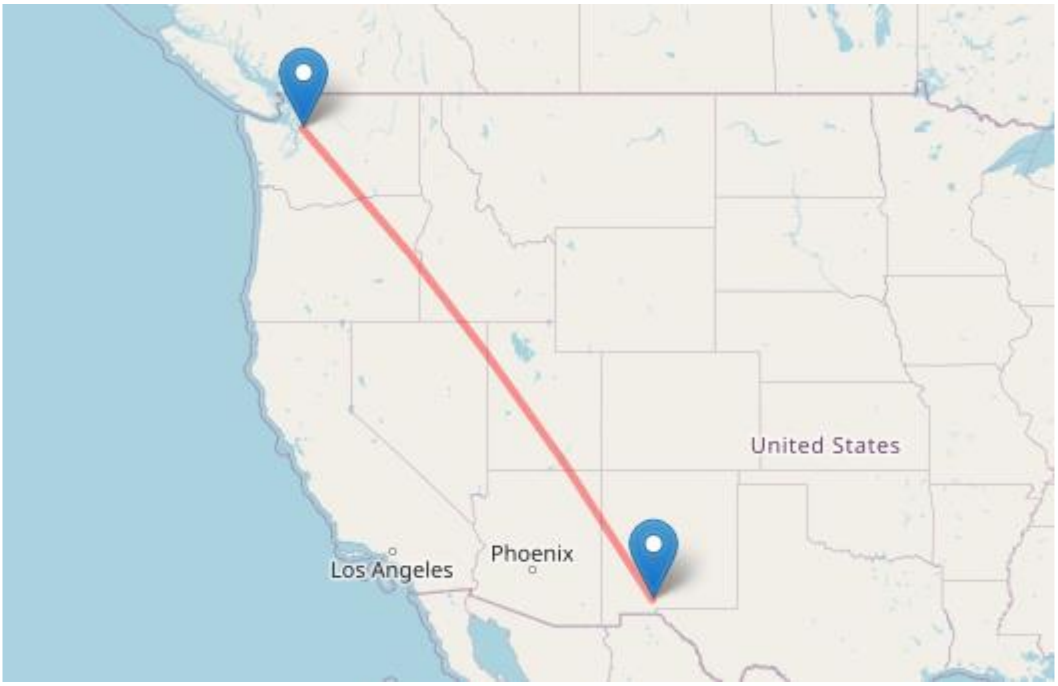
Main characteristics of the network:

Number of nodes:	4941
Number of edges:	6594
Diameter:	46
Average distance between two nodes:	19.99
Is the network acyclic:	NO
Number of power generators:	Unknown

Real life characteristics:



The Western Interconnection is a wide area synchronous grid and one of the two major alternating current (AC) power grids in the North American power transmission grid. The other major wide area synchronous grid is the Eastern Interconnection. The minor interconnections are the Québec Interconnection, the Texas Interconnection, and the Alaska Interconnections.



As we are working only over the United States and not Canada, we can find that the maximum distance between two cities that belong to the western interconnection is 2170 km.

It is important to notice that this distance is not necessarily the one that is covered by the diameter of our graph.

The graph is unweighted, how can this graph model reality?

The 2170 km distance is quite small considering the fact that the velocity with which electricity travels in supply lines is more than enough to cover the 2170 km in less than a minute, this is why we won't consider the fact that the graph is unweighted a problem for us.

This network is a small world network

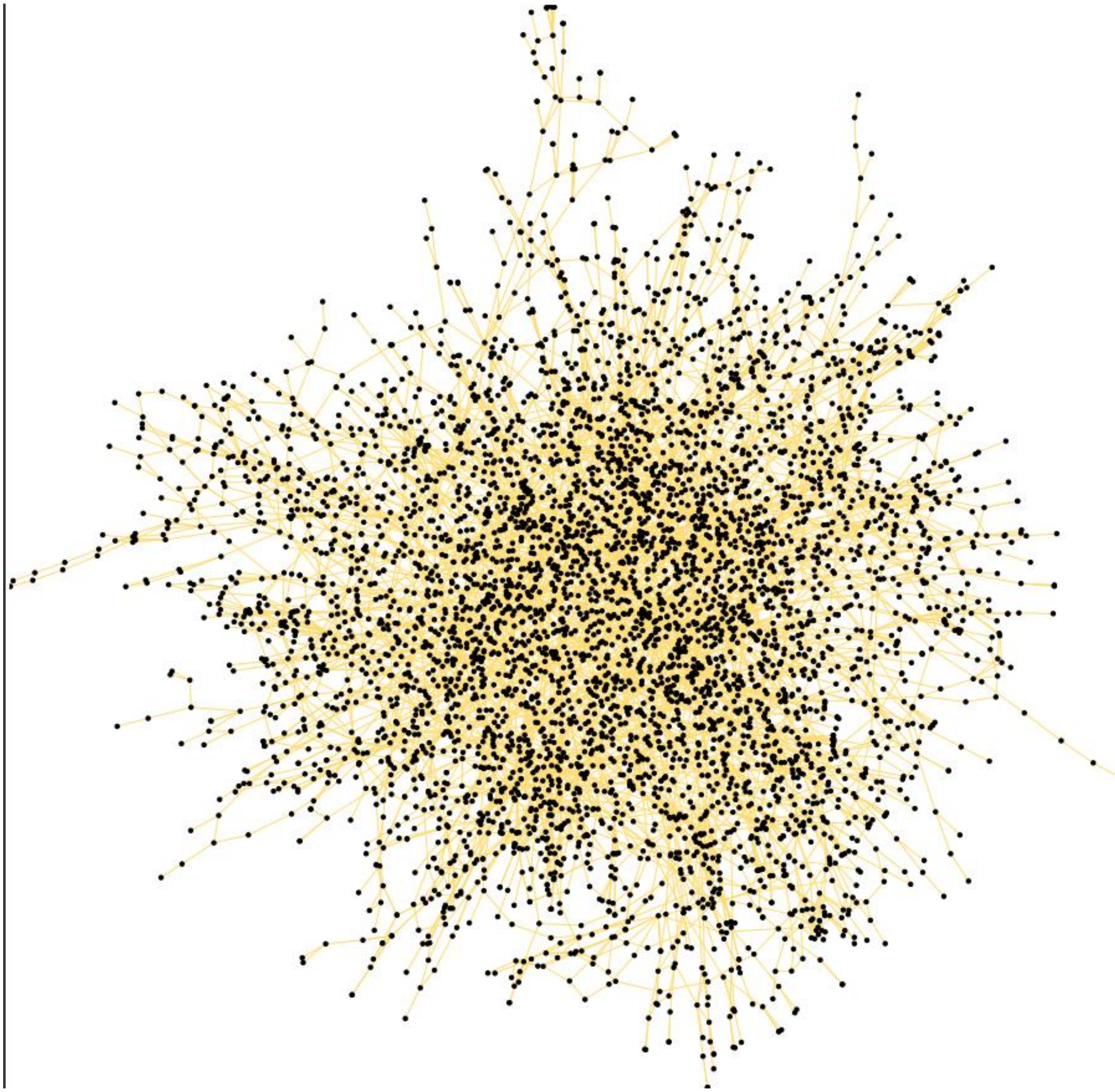
A small-world network is a type of mathematical graph in which most nodes are not neighbors of one another, but the neighbors of any given node are likely to be neighbors of each other and most nodes can be reached from every other node by a small number of steps.

Table 1 Empirical examples of small-world networks

	L_{actual}	L_{random}	C_{actual}	C_{random}
Film actors	3.65	2.99	0.79	0.00027
Power grid	18.7	12.4	0.080	0.005

Table obtained from Watts & Strogatz article.

Initial plot of our graph:

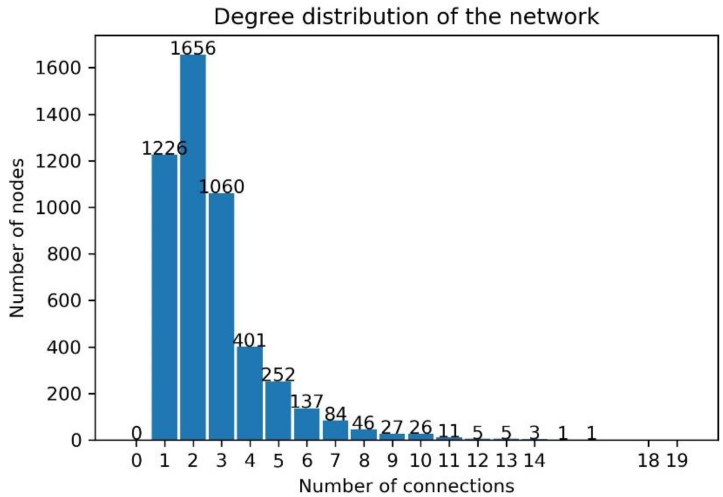


Is not an easy graph to visualize, and even worst when we don’t have idea of the meaning of each node.
This leads to our first question, which will be answered in the following pages.

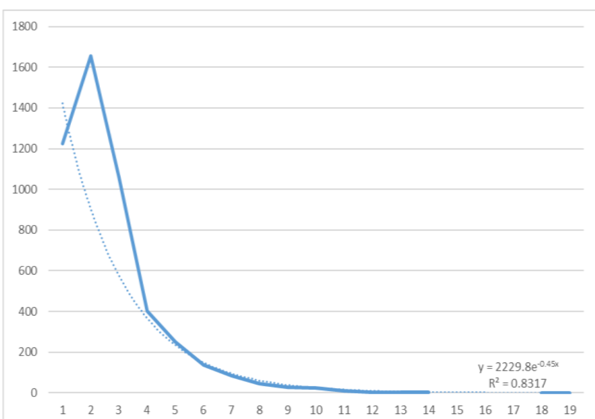
1st question: Which nodes are power generators?

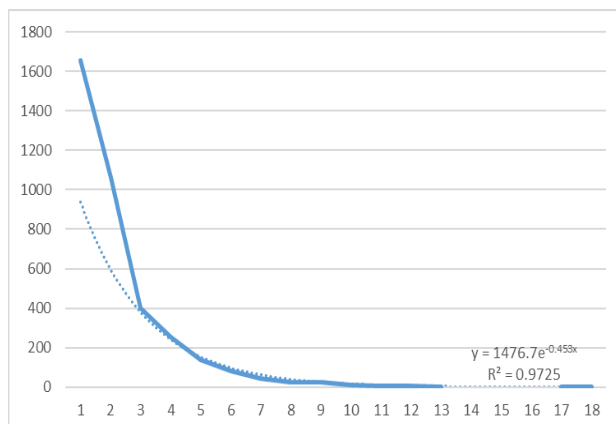
Analysis of the network:

Degree distribution:



Regression of the degree distribution:





We see that R^2 is not quite a good value for our regression, but this is probably given by the number of nodes which have degree equal to one. Once we remove that first value, the R^2 goes up to 0.9725, which tells us that our original curve is very approximate to an exponential form.

Centralities:

In graph theory and network analysis, indicators of centrality assign numbers or rankings to nodes within a graph corresponding to their network position.

What is an important node for our network?

As we are analyzing a Power Grid, a power generator is what we define to be an important node for us, but it is really hard to find out exactly which nodes are power generators, as we don’t have any more information than the graph itself.

In order to find out our important nodes, we will take the best nodes ranked by centrality, and check for intersections with other centralities.

For this network, we will initially consider two important centralities: Degree centrality, and PageRank.

Degree centrality:

Degree centrality is defined as the number of links incident upon a node. This was already calculated and shown with the degree distribution charts.

This centrality has a value for the analysis of this specific network, as we might notice that an important node (Power generator) will tend to be connected to a lot of other nodes.

PageRank:

PageRank centrality is a variant of EigenCentrality designed for ranking web content, using hyperlinks between pages as a measure of importance. It can be used for any kind of network, though.

Each node in a network is assigned a score based on its number of incoming links. For our specific network, this is the same as outcoming nodes, so we might find great information coming from PageRank centrality.

Top 60 nodes by Degree centrality:

2553, 4458, 4345, 3468, 831, 3895, 2585, 2575, 2542, 2382, 2662, 2617, 2439, 2434, 1224, 4395, 4384, 4381, 4373, 4352, 4332, 2282, 1334, 1309, 1005, 490, 4402, 4392, 4361, 4359, 4346, 3838, 3355, 3312, 3128, 2936, 2851, 2800, 2717, 2608, 2586, 2554, 2533, 1460, 1326, 1170, 1166, 1106, 1091, 1050, 1030, 725, 4878, 4391, 4342, 4336, 4199, 3351, 3329, 2959.

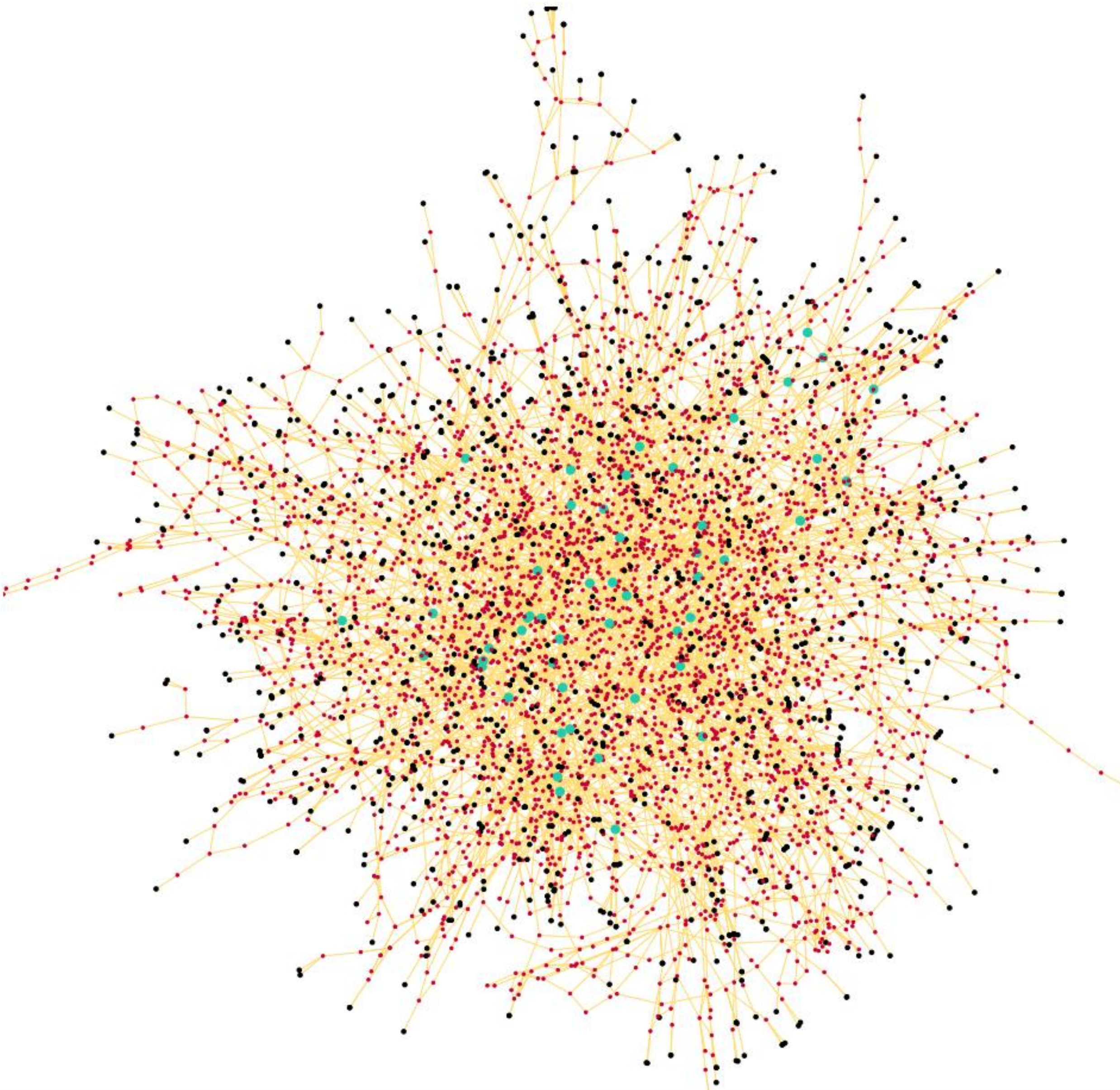
Top 60 nodes by PageRank centrality:

4458, 831, 3468, 2553, 1224, 597, 2382, 2575, 2439, 3895, 3355, 2434, 2542, 2617, 2647, 2282, 1005, 4520, 1030, 1334, 2585, 1326, 4878, 490, 4224, 725, 1554, 1460, 3838, 1106, 2321, 98, 274, 2662, 2249, 2717, 3312, 846, 2586, 2608, 1309, 848, 88, 3979, 3351, 3329, 1166, 3431, 2936, 1091, 129, 2548, 2493, 854, 1098, 2800, 4209, 4199, 803, 3411.

Power generator nodes:

2561, 2434, 2439, 1416, 1166, 2575, 2959, 1170, 3352, 2585, 2586, 2458, 1309, 2596, 680, 2604, 2607, 2608, 2554, 2485, 1334, 2617, 2489, 2235, 1090, 1091, 2883, 3783, 1353, 4172, 1361, 1106, 2522, 4448, 1506, 2533, 2662, 4199, 2918, 4458, 490, 2538, 2542, 3312, 2545, 1140, 3317, 2548, 2553, 4218.

Plot of the network with power generators found:



The green nodes, are where our generators are located, black nodes are transformators (leaf of graph) and the red nodes might as well be transformators or substations.

Having found this graph, now it’s going to be easier to work with the network.

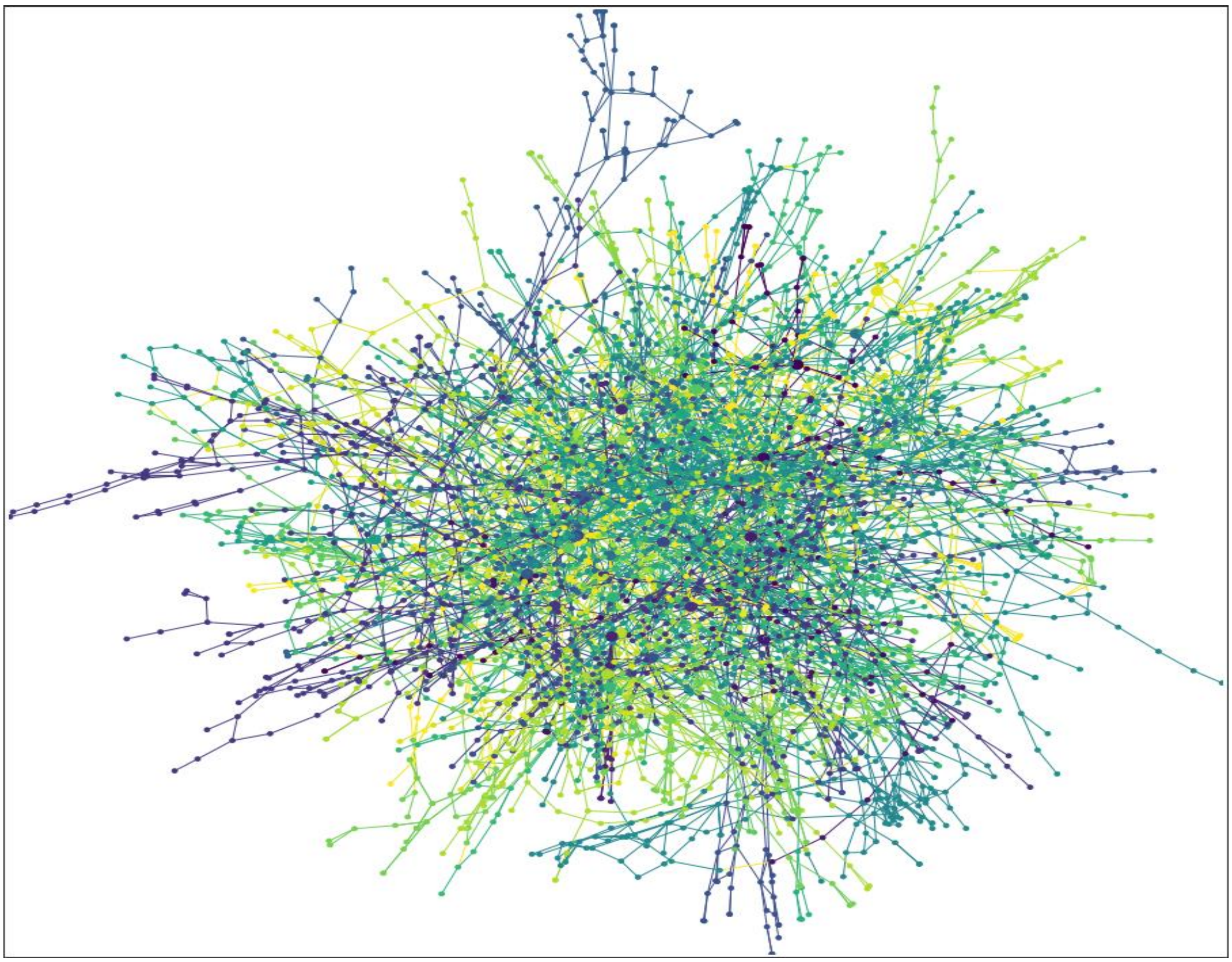
As well our first question has been answered.

Communities in our network:

A community for our network, is a set of nodes which are being powered by one power generator.

We will define: for each node, the power generator that provides electricity is the closest power generator.

After finding this, we will get the following plot:



Where each color, belongs to one community.

As this graph is really difficult to visualize, I plotted each community separately, I will not add them in this document, but still the plots are located in the folder called *communities*.

Communities contain from only 15 nodes, up to 467 nodes.

Sizes of communities:

Size of community of 2282: 245

Size of community of 3329: 210

Size of community of 3355: 44

Size of community of 1091: 391

Size of community of 1030: 213

Size of community of 4458: 318

Size of community of 725: 179

Size of community of 1309: 262

Size of community of 1326: 94

Size of community of 1334: 85

Size of community of 4199: 211

Size of community of 490: 140

Size of community of 2382: 143

Size of community of 831: 68

Size of community of 1005: 72

Size of community of 1166: 117

Size of community of 1106: 396

Size of community of 1224: 69

Size of community of 1460: 87

Size of community of 2617: 210

Size of community of 3312: 467

Size of community of 2542: 95

Size of community of 2586: 110

Size of community of 2434: 68

Size of community of 2439: 24

Size of community of 2575: 15

Size of community of 2717: 108

Size of community of 2553: 79

Size of community of 2608: 84

Size of community of 2936: 34

Size of community of 2585: 18

Size of community of 2662: 32

Size of community of 2800: 21

Size of community of 3468: 48

Size of community of 4878: 84

Size of community of 3351: 16

Size of community of 3838: 51

Size of community of 3895: 33

Notable communities:

Largest communities	Size	Smallest communities	Size
3312	467	2575	15
1106	396	3351	16
1091	391	2585	18
4458	318	2800	21
1309	262	2439	24

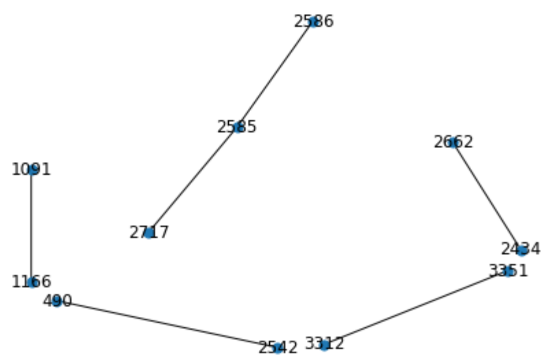
2nd question: Why are there communities of 15 nodes?

This question is difficult to answer, probably a reason for this is that the community is very separate, and only needs a small power generator.

Another reason might be that these communities are actually part of a bigger community that has diverse power generators instead of one large power generator.

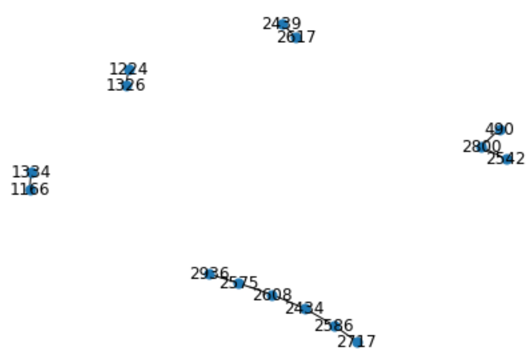
When we calculate the shortest path between all generators, we find that the following list of 2 generators are at a distance of 1 from each other:

2434	2662
2585	2586
2585	2717
1091	1166
490	2542
3312	3351



And the following are at a distance of 2:

2434	2586
2434	2608
2439	2617
2575	2608
2575	2936
1166	1334
2586	2717
1224	1326
490	2800
2542	2800



We check each case separately, and we notice that in many cases, its just two power generators that feed different large communities, and happen to be close to each other, in given case this is a positive scenario, as the failure of one power generator might not have a big impact, as we have a large generator close to sustain the communities.

We also have the case of two generators very close, one feeding a large community, and other feeding a small one, for example: 2439 and 2617, feeding communities of size 24 and 210 respectively. For given case, we have for sure that if the small generator fails, the large one will be enough to feed both communities, but if the big one fails, most probably the small one will be overcharged when rewiring, a possible solution is to disconnect these power generators, or make it a connection that only benefits the small one.

And the final case, of two small generators connected, for example 2575 with community of size 15, and 2936 with community of size 34. In this case I believe its highly recommendable to upgrade one of the generators so it can provide for both, as it doesn't make much sense to keep both active.

3rd question: What happens in case one power generator fails?

We are going to analyze this question under the assumption that a power generator can provide 2 times its original capacity.

If a power generator shuts down, then the network is automatically reconfigured in a way, such that we can only use the power suppliers we already have, and each node will receive electricity from the closest working power plant. If after this reconfiguration, some other power generator gets overloaded, it will also fail, creating a cascade effect in our network.

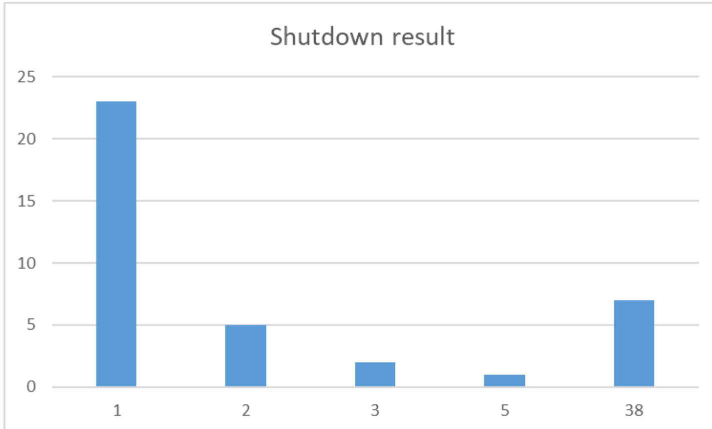
After simulating the failure of every power generator, we get the following result:

When power generators **2617, 1091, 1106, 4199, 4458, 490, 3312** are shutdown, all the network is completely shut down, leaving every node in our network without electricity. It's important to see that 4 out of these 7 are in our top 5 of largest power generators in the network.

When power generators **2434, 1030, 2586, 1309, 2717, 2608, 2936** are shutdown, 1 or 2 other power generators also fail, but doesn't lead to a full shutdown in our network, this is an acceptable situation as all the network is still receiving electricity.

When power generator **2553** is shutdown, causes that 4 other generators fail, which could be a problem when we have to repair those generators, but it's not an immediate problem for all the network.

For all the other power generators, nothing important happened besides the only generator failing.



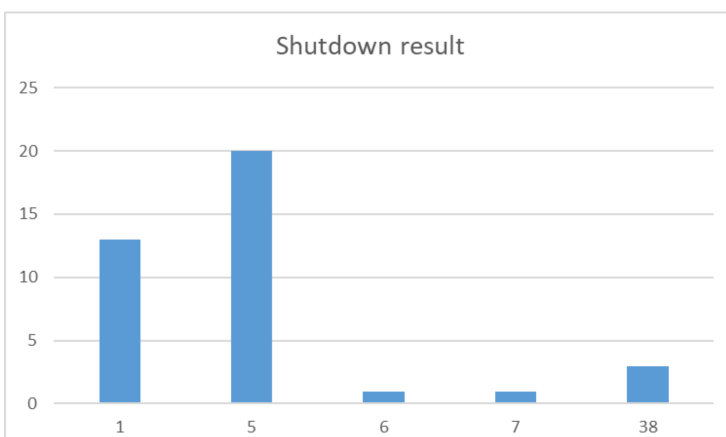
For the power generators that create a major failure, we notice these generators are giving electricity to large communities, above 100 nodes each.

It’s highly recommendable to have an extra power supply for emergencies for all these power generators, but also recommendable to rewire the network in such a way that the failure of these generators doesn’t lead to such a catastrophic result.

A simple and logic rewiring, is to create a clique around all biggest power generators, or at least a cycle, such that they can rely more effectively on each other.

What happens when we create a clique over the power generators?

It’s interesting to see that after we connect all generators as a clique, the possibility of a total shutdown gets lowered, but a shutdown on a single generator tends to cause a small shutdown over more than just itself.



In this case, only generators **1309, 1091, 3312** cause a total shutdown, which makes them even more important for the network, there has to be a bigger plan in case these generators fail.

Questions still necessary to discuss:

- Which edges are necessary to add to this network in order to avoid a total shutdown?
- Where in our network would be really useful to build a power generator?
- Which nodes are dependent on only one power generator?

All the code used for analyzing this network is available in the document ***analysis.ipynb*** together with this document.

References:

Collective dynamics of ‘small-world’ networks, D. J. Watts and S. H. Strogatz, Nature 393, 440-442 (1998).

https://networkx.org/documentation/stable/