United States/North America Power Grid: Network Analysis of Major Blackouts

Mohamed Rakha

Introduction:

Power grids are great examples of modern infrastructure that can be represented as interconnected networks. In my analysis, I am analyzing the US/North America Power Grid, represented as a network with 4941 nodes, representing either generators, load buses, etc as well 6594 edges, which are the transmission lines of power between two electrical components. Some basic network properties include a clustering coefficient of 0.07 and an average shortest path length of 46.98, meaning this network does not have small world properties.

From our preliminary analysis, we can see that because we have a large average shortest path length, one generator tripping, actual hackers, or mechanical distortion along the route can have a bad impact on the rest of the grid. Although power flow and quality engineers have made it so that they can easily isolate the point of failure and load the backup generators/lower load impact, there are possibilities where instances can spiral. The question I hope to answer is: how immune is the current infrastructure to major blackouts?

Throughout this analysis, we are going to look at the network in relation to subnetworks that relate to the the infamous 2003 Northeast[1] and the 2011 Southwest blackouts and show if any potential mishaps in the transmission of power can lead to major power outages in the future.

---

[1] "What Caused the Power Blackout To Spread So Widely and So Fast?" PR Newswire. Genscape. August 15, 2003.

For some light background, in 2003, a software bug in grid operator's alarm system in Ohio resulted in 256 power plants shut off within four hours (due to cascading failure). In 2011, the one line (500 kV) connecting SD and Arizona shut down due to technician error. And although it was reconnected, the phase difference was offset, resulting in line and transformer overloads.

To guide our analysis, we will be making the assumption that power grid networks are not 'rich-club' networks according to Csigi, M., Kőrösi, A., Bíró, J. et al in their paper, *Geometric explanation of the rich-club phenomenon in complex networks*.[2] In their analysis, they calculate edge lengths and how they change with respect to a network's geometry. By analyzing the power grid network as their non-rich network, they conclude that "networks with no rich-clubs [limit]...the length of the connections." What it means to not be a rich-club network is that hubs are NOT densely connected (ie very sparse). Therefore, we can make the conclusion that we should expect a very small degree and betweenness centrality in the network. This is because "the electric current cannot be transferred efficiently (i.e. without huge losses of energy) over large distances without intermediate transformations at middle stations."

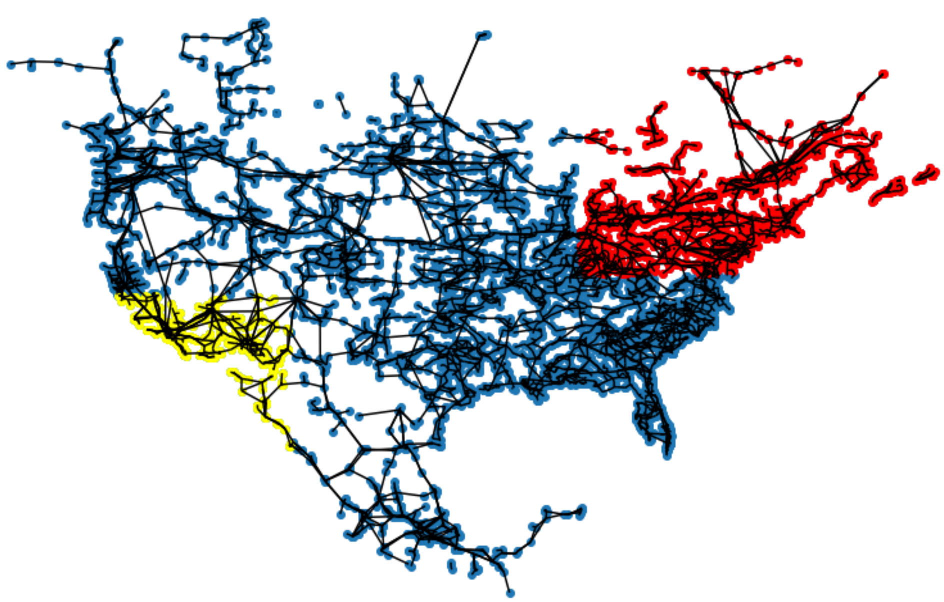

Figure 1- Drawing of Network (Nodes and Edges)

[2] Csigi, M., Kőrösi, A., Bíró, J. et al. Geometric explanation of the rich-club phenomenon in complex networks. Sci Rep 7, 1730 (2017).

Methods and Results:

In order to see whether or not the infrastructure is responsible for the the blackouts, we will be doing a comparative analysis of the power grid network as a whole as well as the subnetworks of the states/areas impacted by the 2003 and 2011 blackouts (In Figure 1, the 2003 impacted area is highlighted in red, and the 2011 impacted area in yellow). I chose to use GeoPandas and NetworkX Python modules to graph and analyze the dataset.

In a power grid network, in order to ensure one mishap (line going offline) does not result in a cascading failure effect, we need to minimize path length and optimize closeness centrality score (where shorter distances are better). This is quite intuitive, as minimizing the path length would mean less lines would be impacted by a single outage across said path because we assume that the disconnection of lines would occur in a successive pattern until that path ends. From Csigi et al, we also know that degree and betweenness centrality will be minimized, but also essentially non-factors.

	Clustering Coefficient	Avg Shortest path length
Non-impacted	0.051	48.00
2003 impacted area	0.053	25.30
2011 impacted area	0.084	14.45

Figure 2 - Small World property analysis of 3 networks

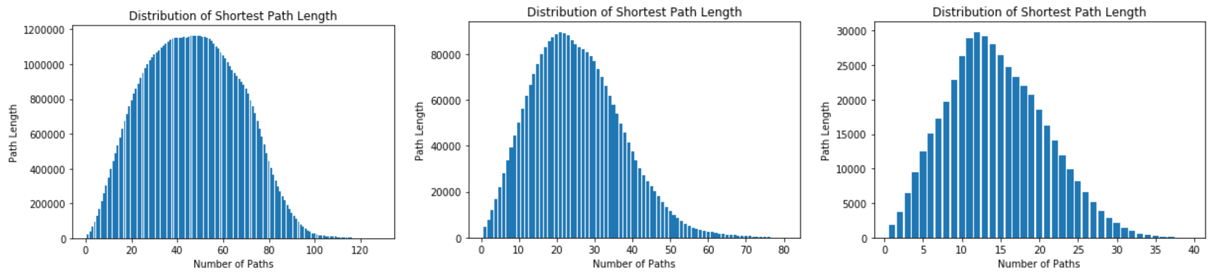

Figure 3 - Distributions of Average path lengths for whole, 2003, and 2011 neworks, respectively

Based on the basic network properties (Figure 2 and 3), we can see that the impacted parts of the network actually had lower average path lengths while the clustering coefficient changed slightly. However, to make a conclusion, we also have to check the closeness centrality scores, which we will take an average of for each network.

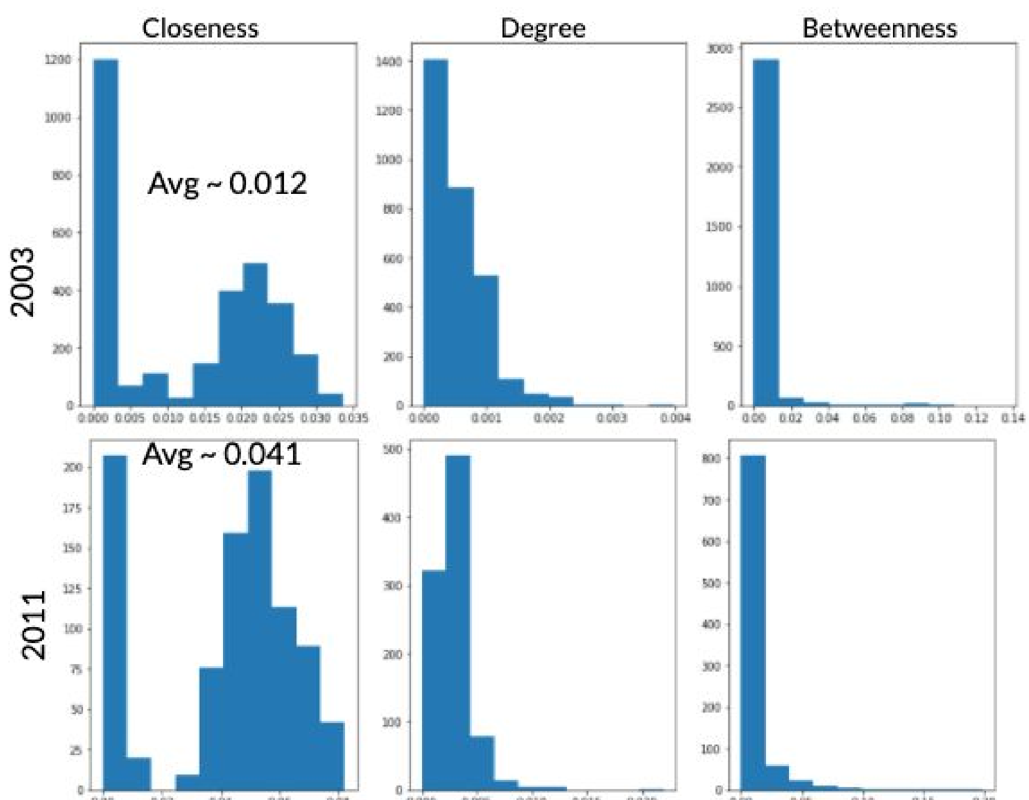

Figure 4 - Distributions of closeness, Degree and Betweenness Centrality for 2003 and 2011 networks

As we can see from Figures 4 and 5, the average closeness centrality scores were 0.022, 0.012, and 0.041 in our total, 2003, and 2011 networks. From the parameters we wanted to optimize (small path length and highest closeness centrality), we see that the 2003 network also had a significantly decreased average path length, but a smaller closeness score as well. However, the 2011 network was the best performing (highest closeness score and smallest path length), and theoretically the least susceptible part of the whole grid.

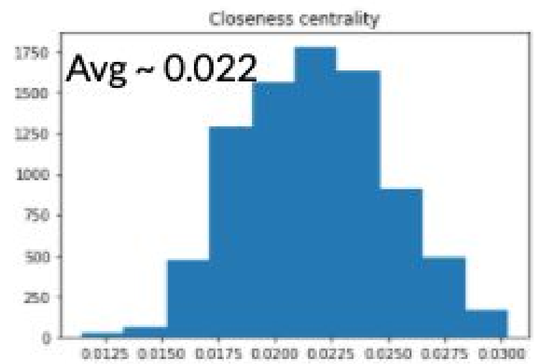

Figure 5 - Distribution of closeness Centrality for whole network

Future Work and Conclusions:

What we can conclude strictly from our computer generated analysis is that major blackouts can't necessarily be attributed to the infrastructure of the power grid network. Larger average path length does not necessarily mean the grid is more or less susceptible. The parameters were way more optimal in the 2011 network, yet it was an area that was hit hard. That being said, because it was a singular line that caused the 2011 blackout due to grid overload, we can say the infrastructure did fail in the sense that adding a node in between that line could have saved the rest of the line. Interestingly enough, more lines were added between Arizona and San Diego to prevent that issue from happening in the future. However we can't really factor that in with the limited data we have.

For future analysis, we can make a network analyzing the same grid and see whether or not material used for generators, transformers, and lines had an impact in the blackout grids. Generally, companies optimize cost over efficiency to build the grid, and perhaps the infrastructure was outdated in the Northeast. We can also take into effect the actual transmission line length and analyze the network as a weighted one and explore the differences. The being said, major blackouts happen every so often, so we shouldn't need to worry too much. We should also remember that there are power flow engineers that monitor the grid and can isolate any faults and manage the load accordingly!