ESD.342: Advanced System Architecture

Regional Power Grid Network Analysis

May 16, 2006

Karen Tapia-Ahumada
JehanZeb Noor
Kate Steel

Advisor: Dr. Daniel Whitney

Executive Summary

The US electric power grid is a complex system with thousands of technical components. This report describes a network analysis performed on one of its regional sub-systems, which includes eight interconnected zones. This research included computation of network metrics such as clustering, degree distribution, betweenness, and path length. From these metrics, we were able to analyze the robustness of the system, based on congestion and dependency on critical nodes. Although we did not have complete information on all other regional power pools, we were able to make some comparisons between this region and both the Western Power Pool and the PJM systems. We also looked at the correlation between the network characteristics and demographics of the region. We found the analysis was useful for understanding the structure of the network, but did not easily translate to socio-economic or financial analysis, for which domain knowledge is necessary. This domain knowledge entails working experience with the power markets, exposure to prices, demand and generation, and network evolution.

Table of Contents

Figures and Tables

1 Introduction

Electric power systems are a collection of generators, transmission and distribution lines, and communication facilities that are physically connected and operated as a single functioning unit. In this system the generators supply a flow of power that is consumed by loads, namely are industries, homes, and businesses. This power flows over a network of connections and is controlled by dispatch centers, which buy and sell electricity to meet the system requirements. Everything about this relationship is subject to pre-approved agreements and regulations. Electric power systems are the very definition of complex engineering systems. The systems have thousands of connected technical components and a management structure that reflects power's status as a social good.

The motivation behind our study was to find how much knowledge, especially related to economic and reliability factors, can be gained for a physical infrastructure network using network analysis. In order for us to achieve this end, we had to build the data up from scratch as we could not find a reliable digitalized source of such information, which is also considered critical infrastructure data. Therefore, the identities of nodes and name of the grid are not revealed in this project report.

2 System description

The focus of this report is a regional power grid in the United States. For confidentiality reasons, we are not able to disclose the location or identify specific attributes of this network. We will therefore, describe the system generally and all analysis will refer to common characteristics of power systems. The exception to this will be our comparison to the Western Power Pool, for which we have published data available. As a note, we are concerned only with the network properties of the technical system. Although we will consider the implications of the architecture on the population it serves, we will not analyze the social or political system of the region.

As a technical system, the regional power grid is composed of standard power system infrastructure (layout shown in Figure 13 of the Appendix 1). This includes: generating stations, loads, buses, and transformers, taps, converters and other connection devices. Figure 14 in Appendix 1 shows the physical layout of the connections. We have designated these structures as *nodes*. The nodes are connected by electrical lines ranging from 69 kV to 345 kV, and also some high voltage DC links. We have designated these connections as *edges*. Overall, the system has more than 1600 nodes and more than 2500 edges. The system has been subdivided regionally into eight interconnected zones, based mostly on geographic boundaries. Later in our analysis, we will consider these zones separately.

2.1 Stimulus, main actors, stakeholders

There are a very wide range of actors and stakeholders in the regional power grid system. To preserve confidentiality, we will frame our discussion in terms of the US power grid as a whole. The various classifications of nodes in the system represent actors, such as generators and loads. Additionally there are management actors, such as the independent system operators (ISOs) and regulators. Finally, every household, industry and business that is connected to the grid is a stakeholder in the system.

2.1.1 Power producers

There are more than 3,170 traditional electric utilities in the US. These utilities can be classified according to ownership, and the system includes utilities which are investor, publicly, cooperatively, and federally owned. In the US, there are 239 investor-owned electric utilities, 2,009 publicly owned electric utilities, 912 consumer-owned rural electric cooperatives, and 10 Federal electric utilities. Although they do not generate electricity, power marketers are also typically considered utilities because they buy and sell power into the grid.

There are approximately 2,110 non-utility power producers in the US. These include energy generators that are exempt under the Energy Policy Act of 1992 (EPACT), independent power producers (IPPs) that sell to the wholesale market, cogeneration facilities with business activities outside of power generation (such as MIT), and other facilities that qualify under the Public Utility Regulatory Act of 1978 (PURPA).

Regarding the energy portfolio for the regional power grid, as shown in Figure 1, about 75% of the energy sources are fossil fuel and the rest carbon free (including nuclear sources).

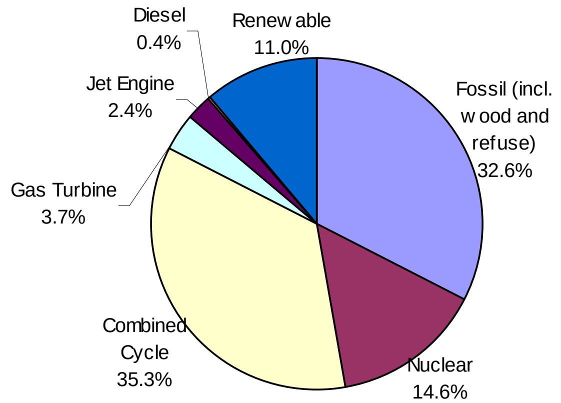


Figure 1: Generation unit by source

2.1.2 Loads

Most of the customers in each zone are residential and municipal. There are also some small businesses. A couple of the zones have significant amount of industry. The load fluctuates between the times of the year, with strong cooling demand in the summer months. There is also a spike in the winter due to heating demand, but this is not as big as the demand for cooling in the summer. Data are available on the load profiles and demand of the customers at a zonal and sub-zonal level. However in order to keep the identity of the power pool in cover and to focus on our network-based analysis, more specific details will not be provided.

In Table 1 we show the principal characteristics of the system we are analyzing at the zonal level.

Zone ID	Population	# Gens	Line (Miles)	Zonal Area (Sq. Mi.)	# Load	Installed MW	Avg. Demand GWh/Month
1	1,321,505	56	1,197	30,862	77	3,391	973
2	1,309,940	36	802	8,968	56	4,205	962
3	623,050	30	403	9,250	29	1,030	507
4	3,510,297	59	1,529	4,844	141	5,405	2,801
5	1,076,189	13	349	1,045	44	2,070	692
6	2,552,265	29	1,047	2,613	85	6,478	1,286
7	3,119,435	38	1 ,262	3,920	107	4,048	1,535
8	727,043	26	789	1,307	78	4,010	2,118

Table 1: Characteristics of Zones 1-8

However, it is to be noted that the zonal demand in MWh is highly correlated to the number of load nodes in the zones and the miles of transmission lines at zonal level. Load demand is relatively strongly correlated with population (although industrial units create demand without increasing population by a lot). Also, load demand has a correlation of 0.63 with amount of MW of generation installed at zonal level. See Table 2 for details.

Correlation Type	Correlation Value
Population-No.ofGens	0.51
Population-LineMiles	0.82
Population-ZonalArea	(0.20)
No.ofGens-LineMiles	0.83
No.ofGens-ZonalArea	0.60
LineMiles-ZonalArea	0.23
No.ofLoads-ZonalArea	(0.09)
No.ofLoads-LineMiles	0.93
No.ofLoads-No.ofGens	0.66
No.ofLoads-Population	0.88
InstallMW-No.ofLoads	0.73
InstallMW-ZonalArea	(0.19)
InstallMW-LineMiles	0.72
InstallMW-No.ofGens	0.35
InstallMW-Population	0.69
Avg.GWh/Month-InstallMW	0.63
Avg.GWh/Month-No.ofLoads	0.88
Avg.GWh/Month-ZonalArea	(0.29)
Avg.GWh/Month-LineMiles	0.71
Avg.GWh/Month-No.ofGens	0.47
Avg.GWh/Month-Population	0.61

Table 2: Correlations between network characteristics

2.1.3 Regulation

The US electric power industry has been transitioning from a highly regulated monopoly to a less regulated competitive industry. When PURPA was enacted in 1978, the generation market was opened to competition. The subsequent EPACT in 1992 further removed constraints to ownership of generation facilities. It also encouraged competition in the business of wholesale electric power sales.

Regulation of the US power grid is handled at the state and federal levels. The state-level public service commissions mainly have jurisdiction over the large investor-owned utilities, although some states do regulate cooperatives and municipal utilities. In general terms, the investor-utilities own more than 75% of the US generation and transmission capacity. They also serve about 75% of US consumers. Intrastate activities power system activities, approvals for most plant and transmission line construction, and retail tariff levels are subject to state regulation. The federal government regulates interstate activities, set wholesale utility rates, license hydroelectric plants, and manages questions of nuclear safety, waste disposal, and environmental issues.

The regional power system that is the subject of this analysis is comprised of all variety of utilities. It has both state and federally regulated activities as it engages in both inter- and intra-state connections. The region contains both nuclear and hydroelectric facilities, which are federally licensed.

2.2 Sources of needs and requirements

The needs of any power system are the demands from the loads on the system. The system needs to supply an equivalent generation capacity, as well as sufficient installed capacity to cover any potential spikes in demand. There are also capacity requirements for the linkages since they must be able to carry the system power flow. As with every design, there is a balance between reliability and over-design. Hence, the system must be able to supply the power requirements over the region efficiently.

For a network representation of the regional power grid, we took into consideration the different attributes for the nodes and links of the system – six and five respectively. We also adopted the following methodology for representing the elements of the network (see Figure 2):

I. For transformers, we added a new node in the matrix for each one

II. For parallel paths, we used an extra node to represent them

III. For taps and junctions, we added an extra node to show the alternative paths

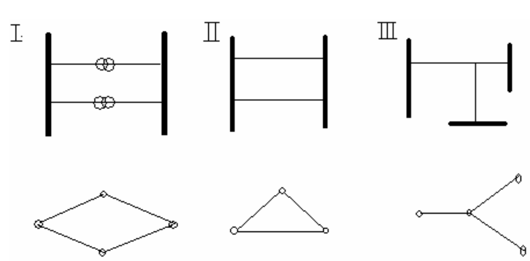


Figure 2: Representation of elements in the network

Using this method, our system representation had 1658 nodes and 2589 links. Figure 3 shows the full network image, with node and edge specifications identified by color as detailed in Table 3.

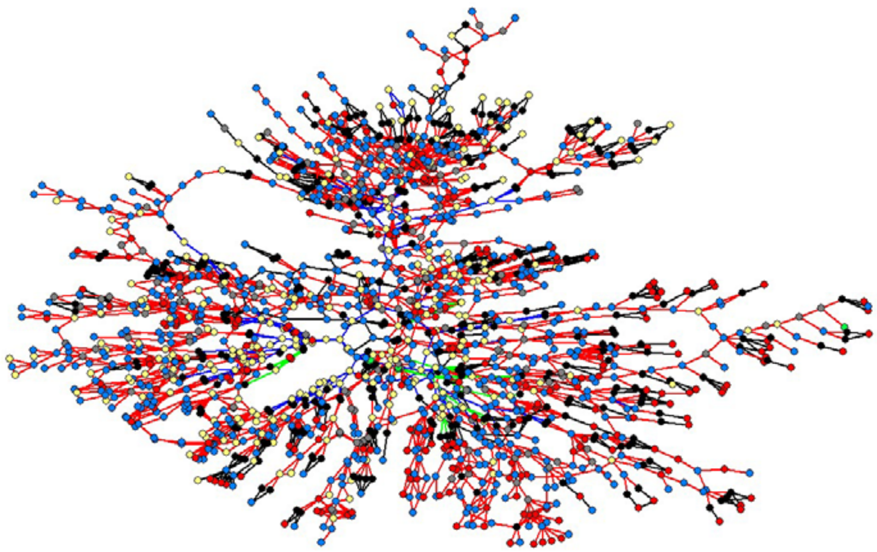


Figure 3: Full network by node and edge specifications

Nodes (type)	
Red	Generators
Blue	Loads
Yellow	Buses
Black	Transformers
Gray	Taps
Green	AC/DC converter
Links (voltage)	
Black	< 115 kV
Red	115 kV
Green	230 kV
Blue	345 kV
Orange	HVDC

Table 3: Color key for network diagram

Although it is difficult to see the range of elements in the overall map, later figures show each zone individually. In these figures the structure of the network can be seen more easily.

2.3 System extent

This regional power grid is spread across several states. As we mentioned previously, within the overall system the Independent System Operator (ISO) defines eight load zones, which represent aggregation of nodes mostly for pricing and trading purposes. In general, characterization of zones follows state geographical boundaries, although there are some zones that have been further divided in smaller areas. These eight zones also coincide with reliability areas that have been considered for the overall region, reflecting operating characteristics and major constraints on transmission system based on past physical conditions. Zones are interconnected with one another, typically only connect through few links. Below is a general description of each of the zones by number.

1. This zone has slightly excess generation as compared to load demand. Most of the load is residential customers, and a good portion of the generation consists of Hydro units. The zone can be termed self-sufficient and rarely has the need to import power. The demographics of the zone are discussed in a previous section of this paper.
2. The load demand here is relatively low and there is rarely any congestion (physical in terms of lack of generation or financial in terms of the locational marginal price). The zone does import some power every now and then in peak heating and cooling months, but overall, appears to be self-sufficient.
3. The load in the zone is also moderate as compared to some congested zones. The zone has a relatively large power plant (nuclear) that exports power to other zones, particularly one of the congested ones. Hence if this generation unit undergoes an outage or the transmission lines are down, it can cause significant congestion in neighboring zones.
4. This is one of the two congested zones, especially one particular part of this zone. This part has a lot of commuters that have high energy demand because the live there, and also a significant amount of office and business buildings. Generation is relatively constrained as compared to demand and a lot of import of power occurs. Consequently, there is congestion especially in peak months and market participants usually use hedging tools to protect their positions. Line outages are critical because

the hamper power import, especially in peak months and especially in the more congested sub-zone (this is not an official zone on its own though).

5. This zone is generation excess and does not generally see any congestion unless there is line maintenance or unit outage. The excess power is exported to neighboring zones that are relatively congested. There are some industrial customers in this zone but demand is still pretty small as compared to the congested zones.
6. This zone has a significant number of generation units and also some industrial customers. However, it is not very highly congested. The critical aspect of this zone is that it has a lot of high voltage lines passing through it to transmit power from high generation to high load zones. Therefore, this zone plays an indirect but significant role in congestion.
7. The role of the zone is similar to zone 6, but the geographic area is slightly larger. The zone also contains the pricing hub that serves as the reference point for locational marginal prices in the rest of the pool.
8. This happens to be the other congested zone in this pool. Population density is high, and generation is relatively constrained. Also, the peaking units are mostly natural gas based which can create huge congestion components and high prices of power. There is a significant amount of import of power in high demand months, with a new transmission line for import into the zone also being constructed.

Figure 4 shows the full network. Colors represent groups of zones as detailed in Table 4.

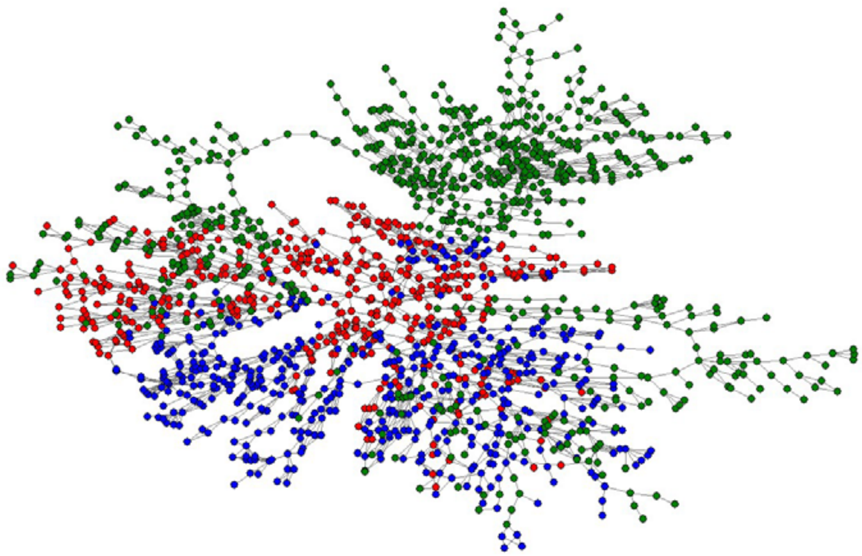


Figure 4: Full network by groups

Groups	
Red	Group 1: Zones 1-3
Blue	Group 2: Zones 4-5
Green	Group 3: Zones 6-8

Table 4: Color key for zone groups in network

The zones were grouped in this manner when we were creating the adjacency matrix and these designations represent the regional layout of the network. Later, the zones are grouped according to similar characteristics, with Group 1 still comprised of Zones 1-3, Group 2 comprised of Zones 5-7, and Group 3 comprised of Zones 4 and 8. Details of these characteristics can be found starting on page 13.

2.4 System mission

The mission of the US power system is to provide power to all customers such that the loss of load is no more than once every ten years, so that well-being of the people and economy can be ensured now and in the future. As quoted by the regional independent system operator, a regional transmission organization fulfills three primary goals:

- Minute-to-minute reliable operation of the bulk electric power system, providing centrally dispatched direction for the generation and flow of electricity across the region's interstate high-voltage transmission lines and thereby ensuring the constant availability of electricity for residents and businesses.

- Development, oversight and fair administration of wholesale electricity marketplace, through which bulk electric power has been bought, sold and traded. These competitive markets provide positive economic and environmental outcomes for consumers and improve the ability of the power system to meet ever-increasing demand efficiently.
- Management of comprehensive bulk electric power system and wholesale markets' planning processes that address electricity needs well into the future.

3 System historical background and evolution

Although we have been asked to discuss the history and evolution of the system, we cannot disclose much of the history because we have to disguise the actual name of the system. Instead of discussing our particular regional grid, we have included some of the history and evolution of the larger US electric power grid.

3.1 Structure of the US power system

The US power system is made up of three major networks (grids), which are then subdivided into regional power pools. The three major networks are:

- the Eastern Interconnected System
- the Western Interconnected System
- the Texas Interconnected System

There are limited high voltage interconnections between the Eastern and Western Interconnected systems, but the Texas System is said to be isolated as it has only DC connections with the rest of the system[1]. There are several international connections, with both the Western and Texas systems connecting to Mexico and the Eastern and Western networks connecting to Canada. There are additional networks in Alaska and

[1] This type of connection allows control over the flow of electricity on those links, contrary to HVAC connections.

Hawaii, which are obviously not connected to the rest of the US power supply. Figure 5 shows the three continental networks.

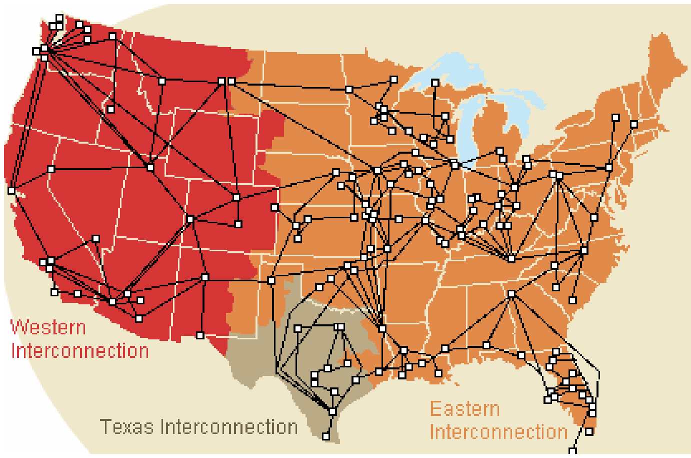


Figure 5: US electric power system

Independent System Operators or Regional Transmission Operator has the role of acting as independent, not-for-profit corporations. To effectively carry out its charge, these entities have a Board of Directors and several hundred employees who have no financial interest or ties to any company doing business in the region's wholesale power market.

3.2 Evolution of the system architectural structure

The US power system has evolved both technically and organizationally over the past 130 years. Technically, it began as an urban system, intended to power Wall Street, homes of the wealthy, and some commercial areas. Over time it was used for street lighting (replacing town gas), for streetcars (replacing horses), and industry (replacing steam). Rural Americans were initially left out of the power system and in 1910 only 13% of US farmers had access to electricity. In the 1930s, under the mandate of the Rural Electrification Administration, farms gained access to electricity and the grid network was extended into rural areas. However, the lines typically ran to pockets of load throughout the countryside, often bypassing towns in between. These “snake lines” zigzagged inefficiently to customers who could afford to pay for the line. In 1935 Morris Cooke described the network as “a haphazard system of finger-like lines, with great pockets of un-served areas between and beyond them.” He further wrote, “Obviously a system planned in advance to cover a given area thoroughly would have been far more satisfactory to the consumers and in the long run to the companies." The zones we are analyzing fit this model, and it is clear from the structure that this is a system which has evolved, as opposed to been designed.

The management of the zones has also evolved. The transmission and distribution, as well as generation, of electricity in these zones used to be regulated. A utility would own much of the lines in one state. Therefore, as per the arrangement of utilities, the zones are mostly defined by utility territory, which usually spans one state. For larger states however, there is more than one utility, such as in northern or southern part of the state. In this case the logical thing to do would be to divide the state into two zones for example. This is what happened when de-regulation occurred and independent system operators took the charge of administering markets in the late nineties. The zones are comprised of state boundaries, with one of the states split on more than one zone. There are transfer lines, mostly high voltage, across these zones or states for import and export of power. Generation and transmission is now separated, with energy marketing companies selling futures contracts of power supply to utilities for serving their load.

4 Size, scale, and network metrics

In this section we will focus on the size, scale and network metrics of the current system. For clarity, we have divided this by zone, although in the final section we discuss the overall power grid architecture. In the following sections, we will discuss Zone 1-3 and Zones 5-7 as groups and Zones 4 and 8 individually. In Appendixes 2 though 4 we show the network model of the nodes and links, using the same color identities as the overall map shows in Figure 3.

4.1 Group 1: Zones 1-3

The three zones that make up Group 1 are the sparsest in terms of nodes and links. They are small networks compared to the other zones and the average path length is smaller than the overall average. Additionally, the edge to node ratio is low for these zones. In these zones there is quite a bit of localized generation. Zone 3, in particular, has several large units which export power to the rest of the grid. Zones 1 and 2 have degree correlations of approximately zero, while Zone 3 has a correlation of about 0.2. This represents a system with highly connected generators. A network map, the degree distributions and basic network metrics can be found in Appendix 2: Network Analysis for Zones 1 to 3.

4.2 Group 2: Zones 5-7

The zones in Group 2 are “intermediate” compared to Groups 1 and 3. They are of mid-range density, both in terms of nodes and population. This group has several high voltage lines, which are the backbone of the regional grid. Unlike Groups 1 and 3, each of the zones in Group 2 has a negative degree correlation. An additional interesting point about the structure of this group is the dependency of Zone 5 on the others. If the interconnections in this group were removed, Zone 5 would be highly isolated. A network map, the degree distributions and basic network metrics can be found in Appendix 3: Network Analysis for Zones 5 to 7

4.3 Group 3: Zones 4 and 8

The Group 3 sub-group is very densely populated and generation-starved. This is particularly true of Zone 4. As a result, the congestion component of the locational marginal price is high. The backbone lines of this Group are smaller, which means there are a number of load nodes coming off of the grid. Zone 8 has quite a bit of industrial and commercial load and high demand, which results in there being two critical nodes in this zone. In this group, the average path length is relatively short and the edge to node ratio is higher. The clustering coefficient is also relatively high and the degree correlation is positive. In Zone 4 it is as high as 0.34. A network map, the degree distributions and basic network metrics can be found in Appendix 4: Network Analysis for Zones 4 and 8.

4.4 Overall system

Table 5 shows the network characteristics of the regional grid as a whole and by individual zone.

Entity	N	M	m/n	<k>	Max k	L	logn/log<k>	Alpha	C	<k>/n	R
WestGrid	4941	6594	1.335	2.67	34	18.990	8.661	No	0.080	0.001	0.004
NewGrid	**1658**	**2589**	**1.562**	**3.117**	**25**	**14.354**	**6.521**	**No**	**0.149**	**0.002**	**0.100**
Zone 1	210	314	1.495	2.952	9	9.514	4.939	No	0.155	0.014	-0.094
Zone 2	139	175	1.259	2.504	10	9.174	5.376	No	0.073	0.018	0.001
Zone 3	92	115	1.250	2.421	7	7.809	5.114	No	0.179	0.026	0.208
Zone 4	393	667	1.697	3.394	25	7.853	4.889	No	0.111	0.009	0.340
Zone 5	88	113	1.284	2.568	9	6.287	4.747	No	0.160	0.029	-0.160
Zone 6	137	288	2.102	2.851	11	9.500	4.696	No	0.231	0.021	-0.090
Zone 7	301	441	1.465	2.93	13	7.902	5.309	No	0.121	0.010	-0.086
Zone 8	221	351	1.588	3.176	15	6.282	4.671	No	0.193	0.014	0.080

Table 5: Network characteristics of the regional grid and zones

Congestion is generally thought of as a financial phenomenon in power markets; however, it is driven by the physical characteristics and temporal or periodic occurrences such as maintenance and planned outages in lines and generating units. Physically, if a unit or line is out, power has to be imported. The higher the degree of the nodes and the power demand, the bigger the effect of the physical outages in the network. Physical congestion translates to a high congestion component in the locational marginal price. This is because the last unit dispatched sets the price of power for the sub-zone, and such a unit is usually an expensive peaker (such as gas turbine). The jump in price between the generating units within the sub-zone and the peaking unit is the financial congestion. Congestion in this report indicates the physical constraints in the system (hence financial by extrapolation, but indirectly – we use congestion in terms of network topology).

On average, there are fewer than two links per node, which means many of them are terminal nodes. The overall clustering coefficient is 0.15 and the average path length is about 14. The Pearson coefficient is positive for the system overall. This is different from technological networks in general. There are constraints on both clustering and connectivity. These constraints are dictated by cost and demographics. In this case, only geographic proximity can define a cluster. The network Betweenness Centralization is equal to 0.33 and the network Closeness Centralization is equal to 0.08.

There are several highly connected nodes, namely 1616 (345kV bus), 1503 (generator), 475 through 478 (generators), 1197 (generator), 1555 (load), 502 (generator), and 687 (load) with nodal degree greater than 10 and in some cases as high as 25. In Appendix 5, we show a visualization of congested zones 4 and 8 in the network, indicated by the yellow lines in Figure 39. The other zones are less congested, have smaller lines and more sparsely distributed nodes. The relative connectivity and congestion across the zones is important when considering the vulnerability and operations of the system. These issues will be addressed in the following section.

5 System effectiveness

As specified by the mission statement, the goal of the US electric power network is to provide the country with a safe, reliable, and cost-effective power supply. From one perspective, the effectiveness of the system relies on its operations, but we also need to be concerned with the absence of operations. We address system effectiveness according to the described mission, but also give consideration to the vulnerability of the system.

5.1 Related to mission statement

The principal concerns from the mission statement relates to the overall operations of the system. Our analysis investigated these concerns with regard to the congestion and resilience of the network. For this analysis, we removed targeted nodes according to their nodal degree k and centrality betweenness Cb. Then we measured the average path length (L) and degree correlation of the system (r). In overall, the analyses showed a diameter ranging from 37 to 74 and a mean geodesic of 14.35 to 25.30.

Figure 6 shows that targeted attacks in the system will result in an increase of the path length. On one hand, the removal of nodes according to their degrees show that the system is particularly vulnerable to this type of attack because: i) there are abrupt increases in L when removing some nodes (particularly when removing nodes 1616 and

1503), ii) the network disconnects in some parts, resulting in a decreased L which is calculated on the remaining connected nodes of the system. On the other hand, the removal of nodes according to their Cb shows that there is a continuous increase of L, reaching values much higher than in the first case. However, the system does not disconnect in subsystems as large as the ones obtained from the previous analysis. This result reiterates the conclusion that power grid is vulnerable to targeted attacks, particularly on those nodes with highest degree.

To test the hypothesis that the system disconnects, we performed a separate analysis on the overall network. Basically, after removing each node we looked for separate components in the network, and we repeated this process after eliminating each highly connected node. The results show that the network separates in big sub areas after targeting specific nodes (for more details see Appendix 6: Effects of nodes 1100 and 1599 on the system).

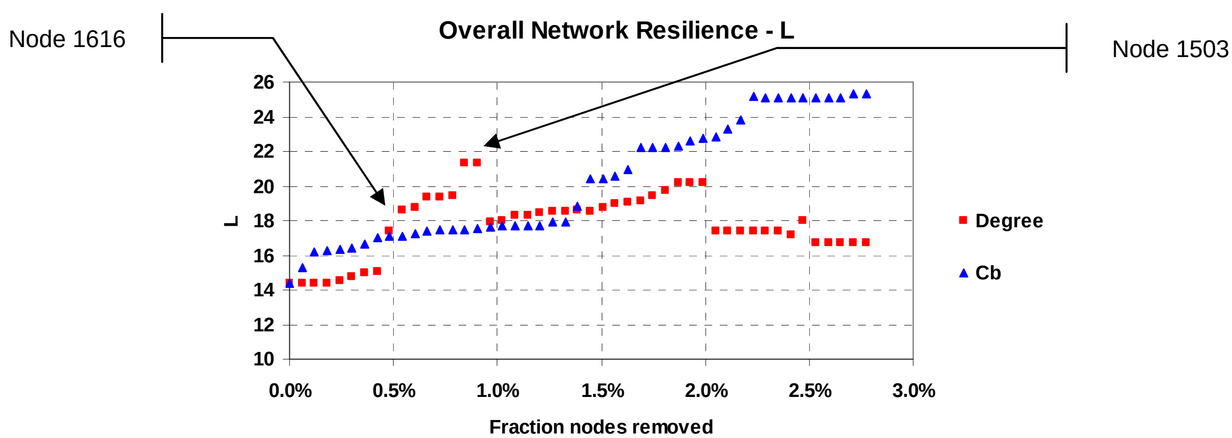


Figure 6: Network resilience - L

In Figure 7, the results show that targeted attacks will impact the topology of the system, and it will reflect changes in the degree correlation of the system. In the first case, the removal of nodes according to their degrees show that the system will change from positive to negative degree correlation, with a tendency to stabilize at -0.06 meaning that the remaining nodes do not present any particular preferential attachment. In the second case, the degree correlation remains positive with some increase not particularly significant. It seems that removal of nodes according to their Cb is not as critical as targeted attacks on high degree nodes.

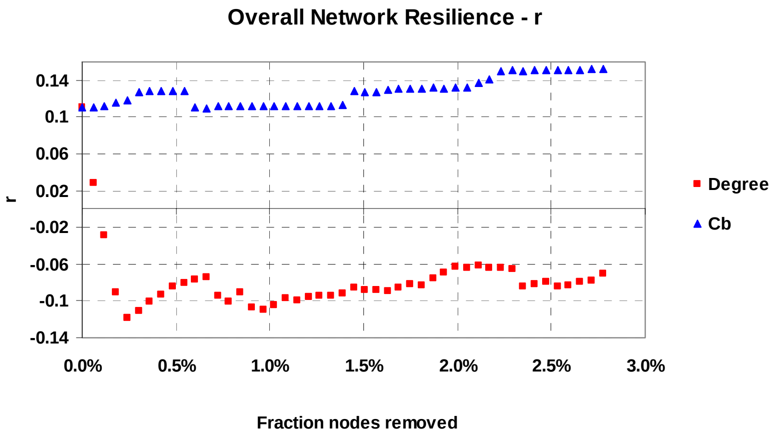


Figure 7: Overall network resilience -r

As first we thought that nodes 1616 (generation units) and 1503 (backbone bus) could be defined as critical nodes. However, we realized that they became important for the system only when considering the cumulative effect of removing the nodes in a particular way. At some point in this process, when several highly connected nodes were eliminated, the system experienced an abrupt increase of the path length (see Figure 6). When removed individually, the impact on the system is the one detailed in Table 6.

Node Removed	Pearson Degree Correlation	Average Path Length
None (whole grid)	0.1018	14.354
Node 1616	0.1117	15.377
Node 1503	0.1007	14.35
Nodes 1616 and 1503	0.1103	15.378

Table 6: Pathlength increase as critical nodes are removed

The effect of losing node 1616 individually is more significant than losing node 1503 (which is almost insignificant). The implications of this are discussed in the section on potential architecture improvements. Finally, we performed a separate analysis for each of the main type of links in the electric system: the 345kV and the 115 kV transmission systems. The calculated basic metrics for this subsystem are the following:

N	180
M	406
L	6.17331
Diameter	14
R	-0.1688
C	0.068

Table 7: Metrics for the 345 kV subsystem

As expected, the clustering coefficient is much less than for the overall power grid since there is much less triangular motifs under this configuration. Also, the topology of the backbone system is reflected in the degree correlation, which it is now negative. The reason for this results rest on the fact that many connections follow the pattern depicted in Figure 8 - several transformers connected to the main buses of the system. In fact, more than 70% of the total nodes of this subsystem are buses and transformers.

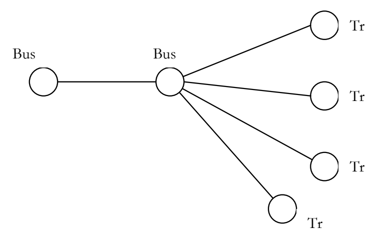


Figure 8: Pattern of connections in the network

In Figure 9, we can see the general network representation for the 345kV transmission system.

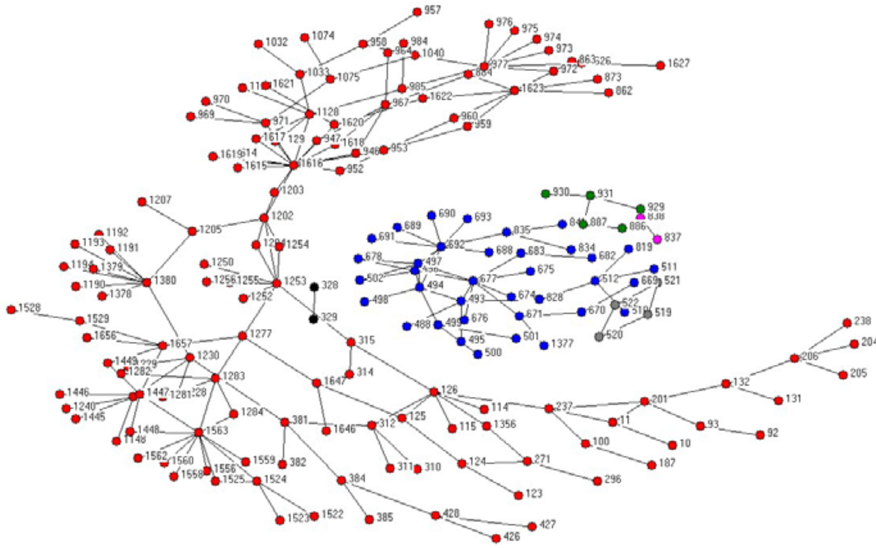


Figure 9: Network representation of the 345 kV grid

In Figure 10 we can see the 115kV transmission system of the power grid. The basic metrics for this subsystem are the following:

N	1318
M	3802
L	18.48238
Diameter	53
R	0.188
C	0.153

Table 8: Metrics of the 115 kV subsystem

In general, these values are close to those metrics calculated for the overall power grid with a slight increase for C and r. The main reason is that, by definition under this configuration, terminal ends of the system are not included[2] (these connections do not show triangular motifs, hence C for the overall network results in a lower value). Regarding the Pearson coefficient, there are many connections that occur between nodes with similar degree and multiple connections between buses with comparable degrees. In Figure 10 we can see general network representation for the 115kV transmission system.

---

[2] Terminal ends are at lower voltage level (less than 69kV), so they are not included for this particular analysis where we are only representing 345kV lines. For the overall analysis of the power grid, terminal ends are included since we analyzing the whole structure of the network.

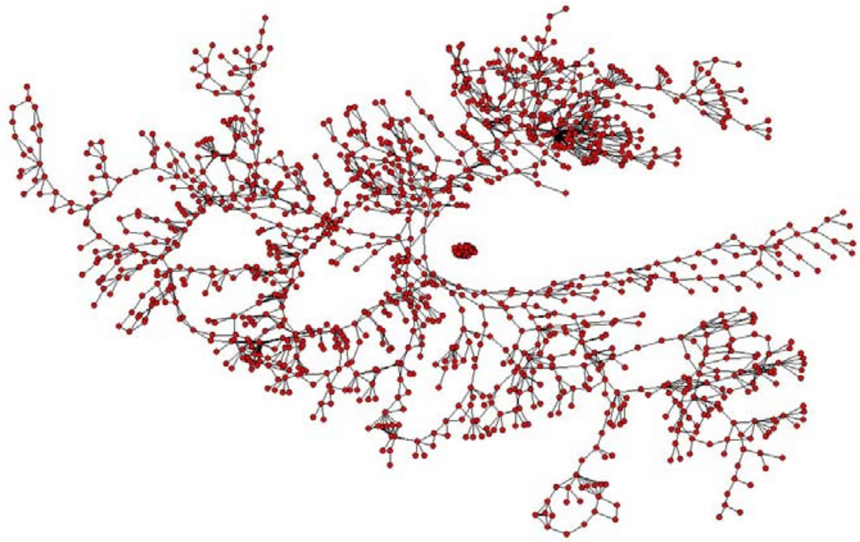


Figure 10: 115 kV transmission system

Then, if we visualized each of these subsystems according to their betweenness centrality we can see the criticality of particular nodes for that particular system. For simplicity, we show in Figure 11 the 345kV system (see Appendix 7 for the 115kV transmission system). From this image we can conclude that nodes 1616, 1202, 1253, 1277, and 126 are critical for the subsystem. Potential problems in these nodes could cause the system to disconnect in big sub areas, as they are common elements for several paths within the system (creating parallel path could be a potential area of improvement for the system).

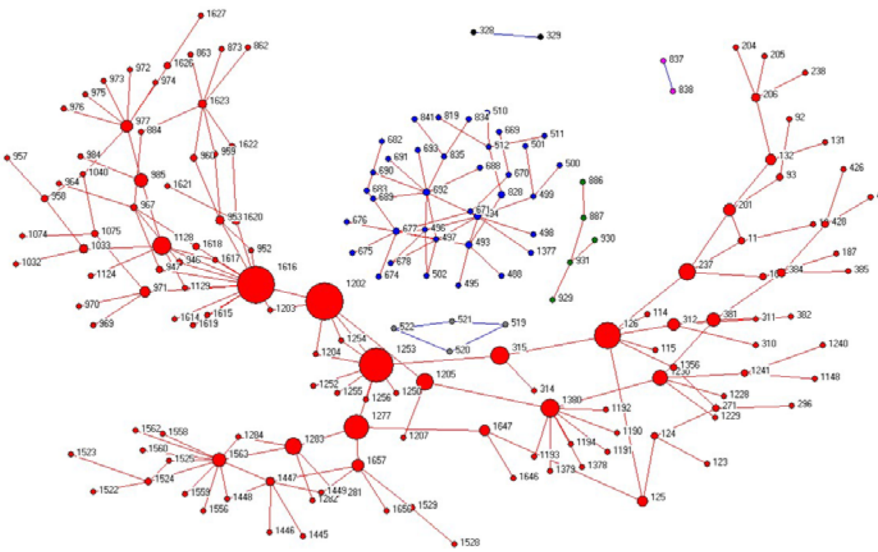


Figure 11: 345 kV transmission system, organized by centrality

Finally, we analyzed the system's hierarchy. Hierarchy is evident from the structure of this adjacency matrix and *betweenness*. In Figure 45 of the Appendix 8 we show a representation of the network. In terms of the adjacency matrix, a quick design structure matrix based partition returns the original matrix, showing the existence of precedence of nodes, which for a physical energy flow system means hierarchy. That is, power cannot reach the nodes lower in the order before flowing through the higher level nodes. The idea of betweenness can similarly be applied. Each block below the diagonal pertains to a zone in this power pool. The off-diagonal entries on the upper right of the graph show inter-zonal and high voltage links.

5.2 Related to system characteristics and –ilities

Post-September 11th one of the major concerns of the US government is the vulnerability of critical infrastructure. The US electric power grid is one of the most critical systems for the functioning of the country's government, businesses, financial markets, and transportation systems. In fact, many of the other critical infrastructures, such as transport, oil and gas, or computer networking, are dependent on the electric power system. As a result, one of the most important characteristics of the system is its robustness in the face of attack.

The present architecture of the electric power system is a highly integrated centralized system. However, it is transitioning to a more competitive system, with consequences for the robustness. As new players enter the market, existing utilities must offer open access to the transmission system. The diversification of utilities involved in generation, transmission, distribution, and the power markets should have economic benefits, but with this diversity comes the risk of opening up the system to players with ulterior motives. Main classes of vulnerabilities in the system are physical and informational.

Physical destruction of pieces of the system is still the greatest potential threat. As our analysis showed, there are nodes which are critical to the network because of their connectivity and load. It is for this reason that the data sets for the power pools has been taken out of public domain. Through our simple analysis, we have identified potential risks in terms of vulnerable nodes. Electric power control systems are the other potential vulnerability. Although recent legislation has increased law enforcement's authority over the system, there are still not effective means of detecting or reporting problems. Of the information systems, substations represent the greatest security vulnerability because many of the automated monitoring devices are poorly protected.

Our analysis reveals some of the vulnerabilities in the system. Removing highly connected nodes has an effect on the average path length in the system. Increasing path lengths means more congestion in the system and was one of the reasons behind the US power outage in August 2003. One example of this type of critical node is Node 687, which is a major generator in the most congested zone. If we were to remove this node it would lead to an increase in the average path length across the system as the power flow must reroute over a less efficient system of paths. If the longer paths could not carry the load this could lead to system failure. Another example, it is the cumulative effect of removing specific nodes that could cause the system to disconnect and form islands within the major system. We will discuss more about the specific network analysis metrics in the next section.

5.3 Related to metrics derived from network analysis or others

We compared our regional power pool metrics with those of the Western Power Grid and a random network generated[3] using the same structure as the grid. The randomly generated network had an average path length of 15 and has a clustering coefficient on the order of 10^{-3}. Table 5 shows the comparison of our zones to this network. The same table shows our regional power pool compared to the Western Power Grid. We found from this analysis that the power grid is a small-world model as average path length are similar and clustering coefficient is much higher. Additionally, we were able to verify our results by the similarity to the Western Power Grid.

Of the metrics calculated, only the clustering coefficient and Pearson degree correlation are much higher. There are structural dissimilarities between the Western and understudy power grids that drive the fundamental difference in these two metrics. Potential reasons include more concentrated load and generation pockets that enhance the degree correlation. Path length is longer because the grid is not as integrated across zones as the western grid appears to be. The results on motifs for the grid understudy are similar to western grid[4]. An analysis of the 3 and 4 nodal motifs using the mFinder software available online shows following results:

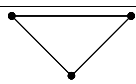
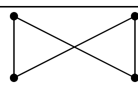
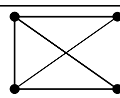
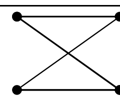
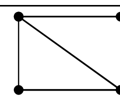
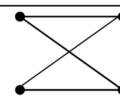
Motif						
Occurrence	450	977	340	632	56	583

Table 9: Motifs in overall power grid

On the other hand, PJM results (for major component of grid) are as follows: diameter = 49, Nodes = 3916, Links = 8446, average nodal degree = 2.2, Pearson Degree

[3] The random network was generated using the Erdos-Renyi algorithm in Pajek software which captures the number of nodes and average nodal degree of the underlying network, the power grid in this case

[4] Motifs on the western grid were presented in Shah's paper for the ESD.342 spring 2006 course

Correlation: -0.07, Average Path Length: 15.389, Clustering coefficient = 0.035. Degree Distribution (Figure 12):

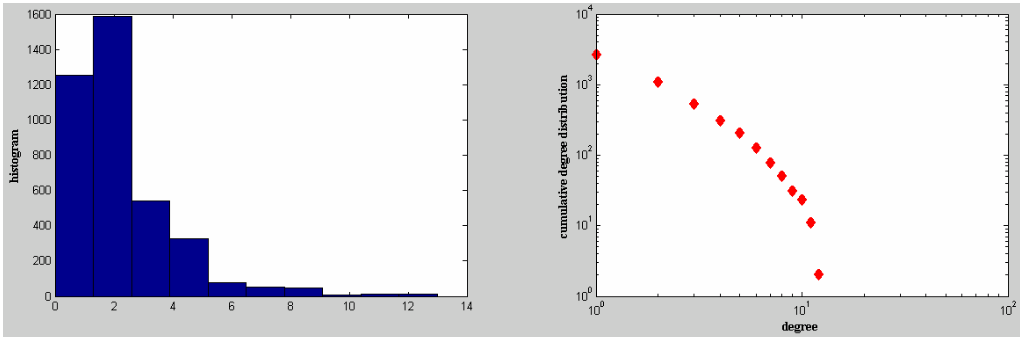


Figure 12: Degree distribution and cumulative degree distribution for PJM

For a random graph generated with the same parameters, the avg. path length is 10.225 and the clustering coefficient is ~ 0. So, we can declare this network small-world. However, it is not scale-free as shown by the second graph. Particular sections of the degree distribution in the middle are scale free. This is what we found for the grid understudy as well.

6 Potential architectural improvements

Any architectural improvements to the system would need to increase the reliability, robustness, and flexibility of the system. To this end, potential improvements include:

- Eliminating overly-critical nodes by balancing the system better
- Increasing the number of pathways
- Creating a more decentralized system to reduce congestion

From our analysis we found there were several highly critical nodes, which had the potential to cause system failure if removed. This makes the system vulnerable to attack or to simple outages. In order to reduce the impact of these nodes the system needs to spread the load across a greater number of nodes and reduce the connectivity of these critical nodes. In addition to creating more nodes, we could also reduce congestion by increasing the number of pathways, as we saw from the 345kV system representation in section 5.1. In doing so, we create more potential routes for the power to flow through, should part of the system fail. Both of these options would increase the reliability but also increase the complexity by adding to the number of components. There is likely a trade-off between robustness and efficiency that we would need to explore more before concretely saying that more nodes and paths would be beneficial overall. Another option to investigate for reducing congestion is to make the system more decentralized. However, this would be difficult in the region of focus, since it is a densely populated part of the country.

7 Reflections

The breadth and depth with which we have performed and presented our network analysis is an exciting addition to the science of networks. While this is one contribution, there are many areas for further exploration, which include but are not limited to:

- incorporation of flow analysis with the network topology
    - studying power flow in normal operation to identify critical nodes
    - studying power flow in constrained operation to understand congestion
    - determining optimality of power dispatch in relation to generation stack
- impact and feasibility of recent additions to network, including 345 kV lines
- layering in the network – studying the subsets of certain node and link types

With a rich set of data and methods already developed, future researchers would be in good shape and would do well to continue the work initiated with this report. In fact, one of the most challenging aspects of this project was to devise a complete and accurate adjacency matrix to capture all types of links, of all voltage denominations, and all types of network components or nodes. Having software tools to check for connectivity was helpful in removing errors - for instance, elements of power grid cannot be disconnected.

We also found that a lot of rich data are available on the internet. A closer study of the locational marginal prices and congestion with respect to physical network topology might reveal more insights than the ones discussed in this paper. Overall, the paper was a learning experience and stimulated interest in the field of network architecture.

7.1 System analogies

Our method of analysis could be applied to a range of systems, particularly other infrastructure systems. It would be very interesting to look at the remaining power systems in the country and compare across regions. These networks could also then be integrated into a full US power system network. Looking at the US as a whole would show more clearly where the vulnerabilities in the system are. It would also be interesting to compare the electricity distribution network with other service delivery/removal infrastructure such as telecommunications, sanitation, oil, and natural gas. However, we are unlikely to have access to this data for the US.

If data were available, it would also be very interesting to compare the US system to power systems in other countries. The US system grew up in a very haphazard way, as was discussed in the section on the history of the power system. In countries where electricity came later, the system might show other hierarchical structure. In class we saw this was the case for subway systems, and it may well be true for power systems.

It is difficult to tell if this analysis would be useful for looking at non-technical networks. It seems that the true insight of network analysis cannot be seen until the calculations are complete so it is best to simply experiment with a range of networks and see what the results are. One potentially similar system would be the human circulatory system. This network has the same characteristic of “flow” through a system of “pipes”. In the circulatory system there is only one generator (the heart) and a large number of loads. A comparison of these systems could yield interesting results.

7.2 Key learning

The principal lessons from this project were:

- Difficulty in obtaining infrastructure data
- Importance of careful matrix building
- Power of software analysis

The time spent in the first two tasks of this project (data collection and matrix building) was out of balance with the time spent analyzing the data. The data was difficult to obtain due to the potential security risk of having detailed information on critical infrastructure. We also spent a large portion of the project time working on building our data set, when there was an easier way we could have built the adjacency matrix. More time spent on up-front planning might have prevented this time expenditure. Once we had our matrix, however, we were able to use UCINET, Matlab, and PAJEK to analyze the data and get results. These proved to be very powerful tools for determining the characteristics of the network, although we needed to be careful on the format of the input data each of the software accepted.

The Role of Domain Knowledge in Network Results and Analyses

We were able to learn a great deal using network theory on the power grid, but two questions arose: how much significant information did we actually learn; and what role does domain-specific knowledge play in our study? Concisely, we learned about the critical points in the physical network, the distribution of generation, the reliability of the network and various metrics that reflect the topology of the system. This information is deemed to be useful in identifying infrastructural improvements in the network such as construction of new lines or more generation. The authors expect that, in fact, the NERC, ISO and FERC look at such network results combined with power flow model analyses (which is beyond the scope of this paper) to decide whether changes in the physical network should be made to meet system goals.

As per the role of domain knowledge, it seems critical in understanding the financial workings of the network. For instance, identifying critical nodes and bottlenecks is not sufficient to predict congestion in locational marginal prices. At best, such identification just points out the nodes to monitor more closely, particularly in peak demand. The user of this network analysis information needs to have seen trends in the trading markets, understand the re-routing methodologies of power, understand the markets of underlying fuels for generating units and incorporate other factors such as weather forecasts in actually predicting congestion. This, perhaps to some extent, explains that while engineers have developed the physical infrastructure and reliability mechanisms, it is economists who have shaped the intrinsic working of the financial models that the independent system operator and market participants adhere to. Domain knowledge seems more critical in market aspects, while network-specific knowledge seems to be more sufficient for studying and improving the physical reliability and flexibility of the network, especially in the case of targeted node failure.

Finally, the authors would like to thank ESD.342 classmates, instructors, and in particular Dr. Daniel Whitney for his continued guidance, support and encouragement. Without this, the project would be far less enriched than it is now.

Appendixes

Appendix 1: General System Description

General figures and diagrams

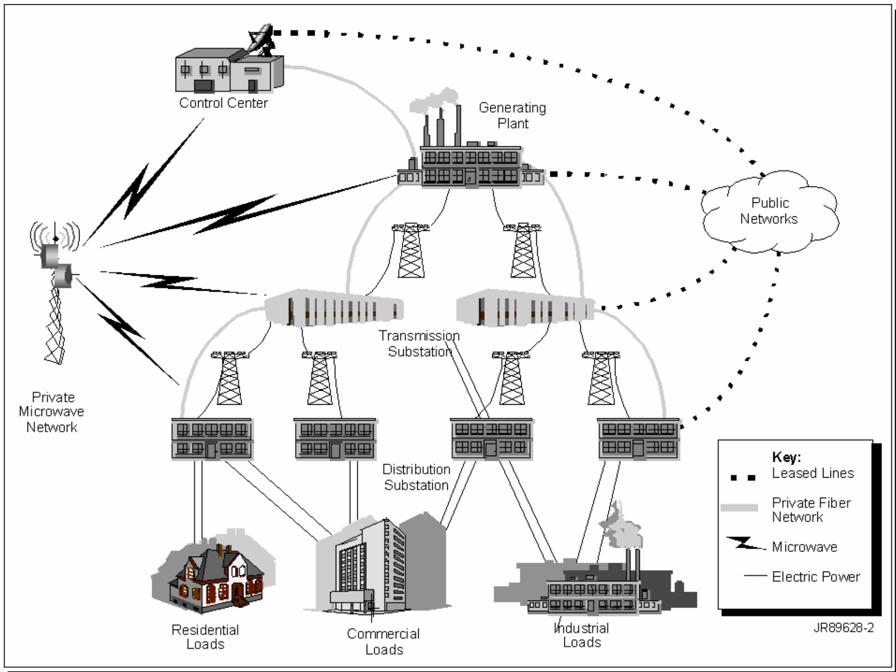


Figure 13: Physical representation of the network

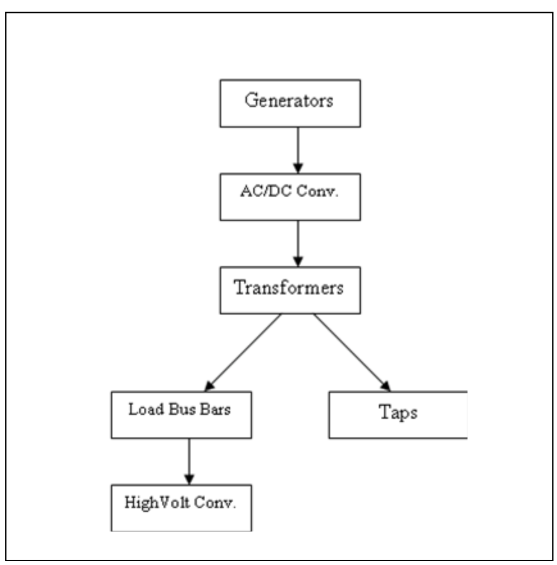


Figure 14: Classification of nodes in the regional power network

Appendix 2: Network Analysis for Zones 1 to 3

Zone 1

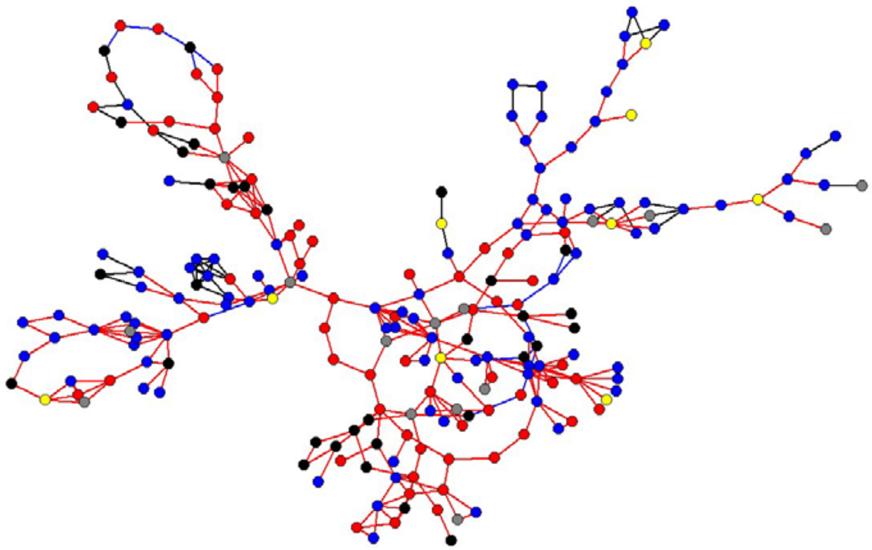


Figure 15: Zone 1 Network

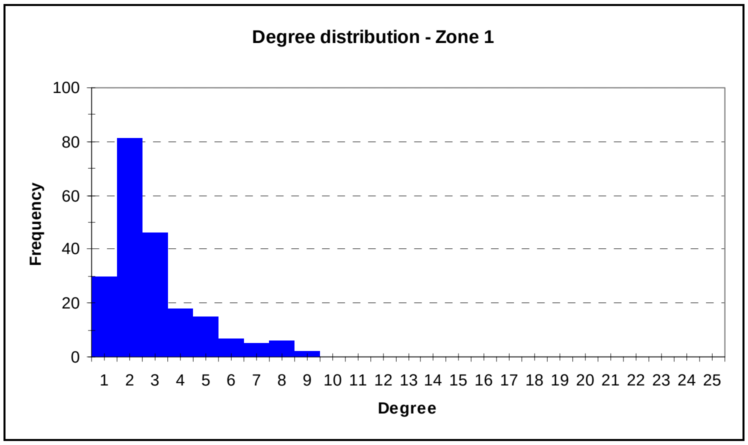


Figure 16: Degree distribution in Zone 1

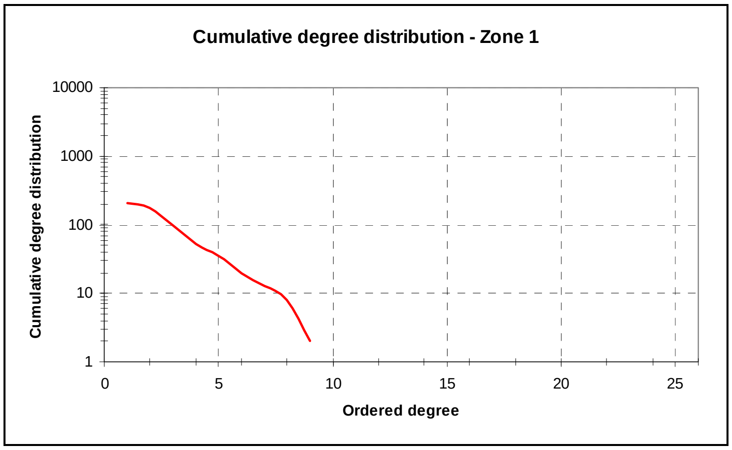


Figure 17: Cumulative degree distribution in Zone 1

	N	M	m/n	<k>	Max k	L	logn/log<k>	Alpha	C	<k>/n	R
Zone 1	210	314	1.495	2.952	9	9.514	4.939	No	0.155	0.014	-.094

Table 10: Network characteristics in Zone 1

Zone 2

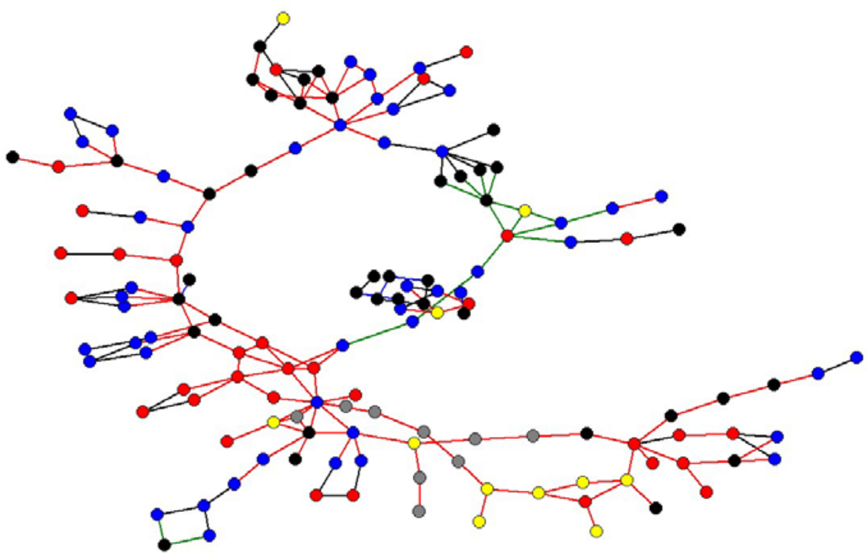


Figure 18: Zone 2 network

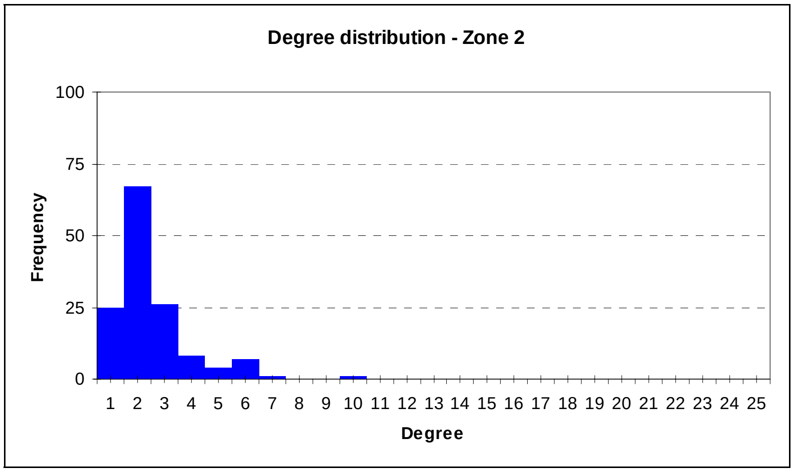


Figure 19: Degree distribution in Zone 2

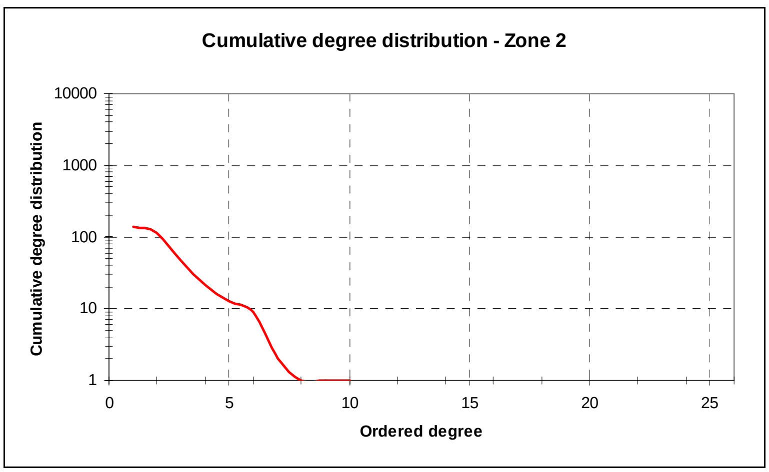


Figure 20: Cumulative degree distribution in Zone 2

	N	M	m/n	<k>	Max k	L	logn/log<k>	Alpha	C	<k>/n	R
Zone 2	139	175	1.259	2.504	10	9.174	5.376	No	0.073	0.018	.001

Table 11: Network characteristics of Zone 2

Zone 3

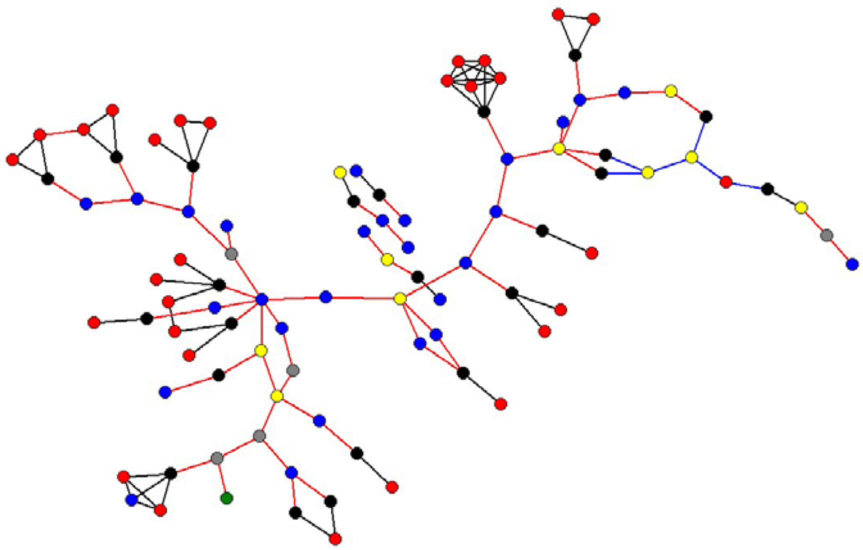


Figure 21: Zone 3 network

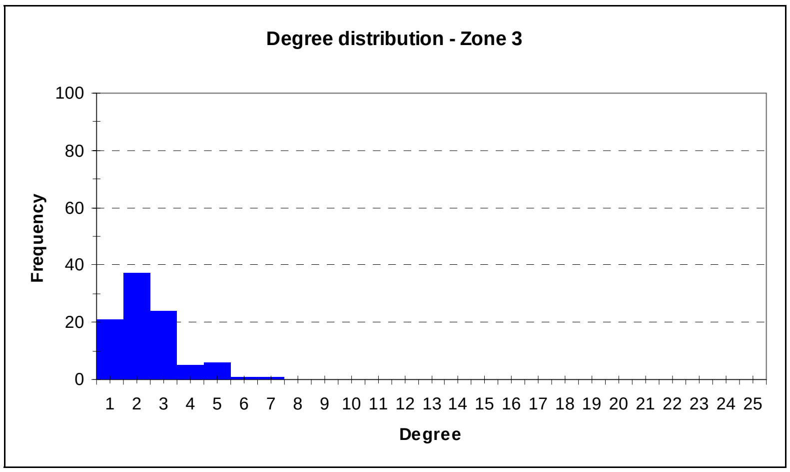


Figure 22: Degree distribution of Zone 3

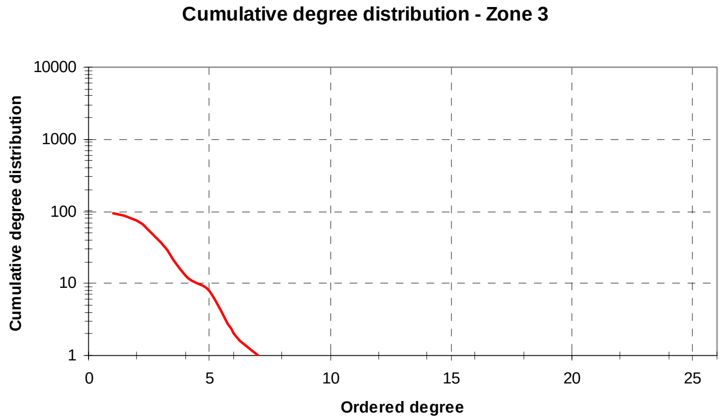


Figure 23: Cumulative degree distribution of Zone 3

	N	M	m/n	<k>	Max k	L	logn/log<k>	Alpha	C	<k>/n	R
Zone 3	92	115	1.250	2.421	7	7.809	5.114	No	0.179	0.026	.208

Table 12: Network characteristics of Zone 3

Appendix 3: Network Analysis for Zones 5 to 7

Zone 5

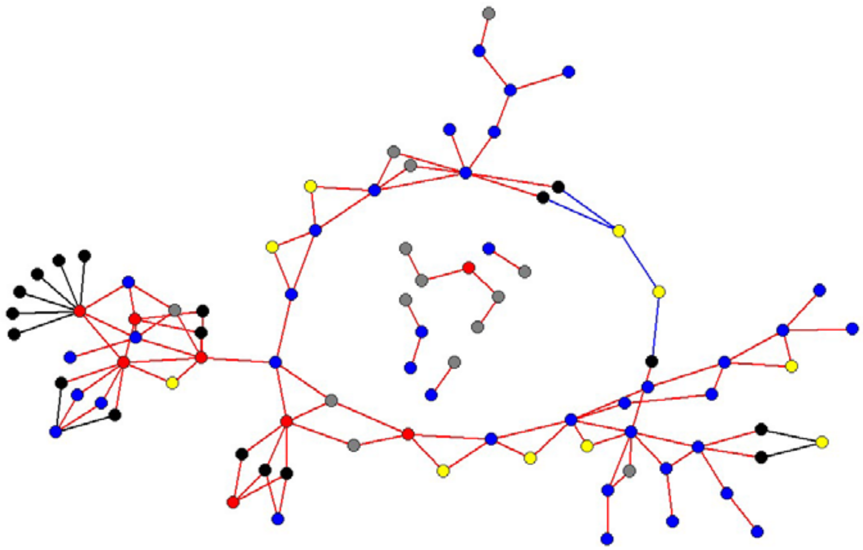


Figure 24: Zone 5 network

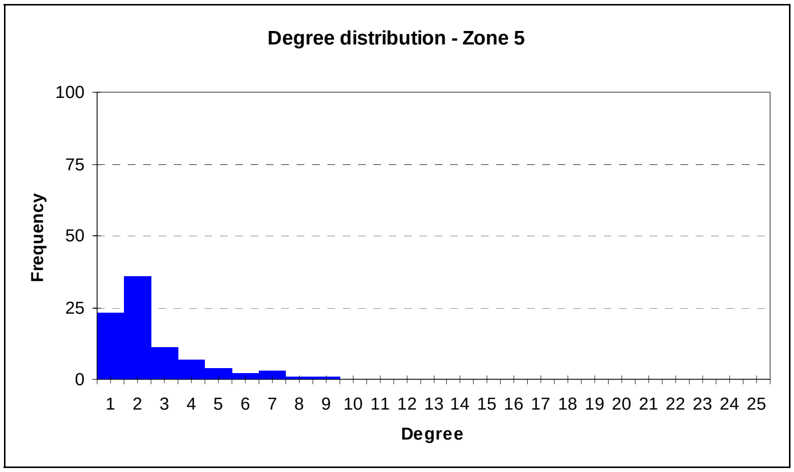


Figure 25: Degree distribution for Zone 5

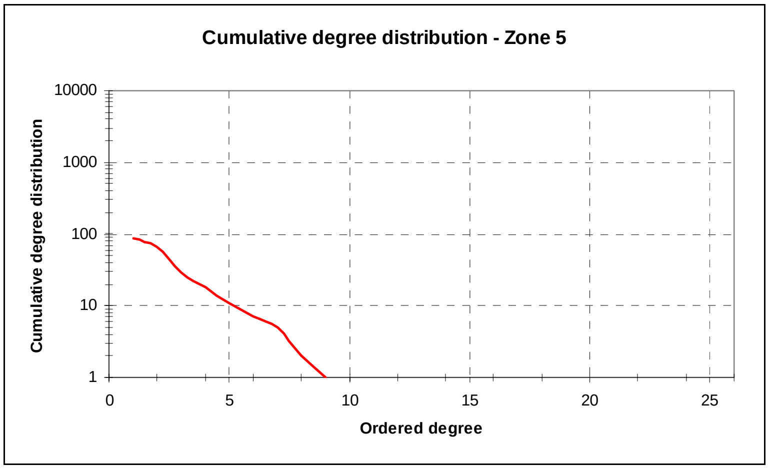


Figure 26: Cumulative degree distribution in Zone 5

	N	M	m/n	<k>	Max k	L	logn/log<k>	Alpha	C	<k>/n	R
Zone 5	88	113	1.284	2.568	9	6.287	4.747	No	0.160	0.029	-0.16

Table 13: Network characteristics of Zone 5

Zone 6

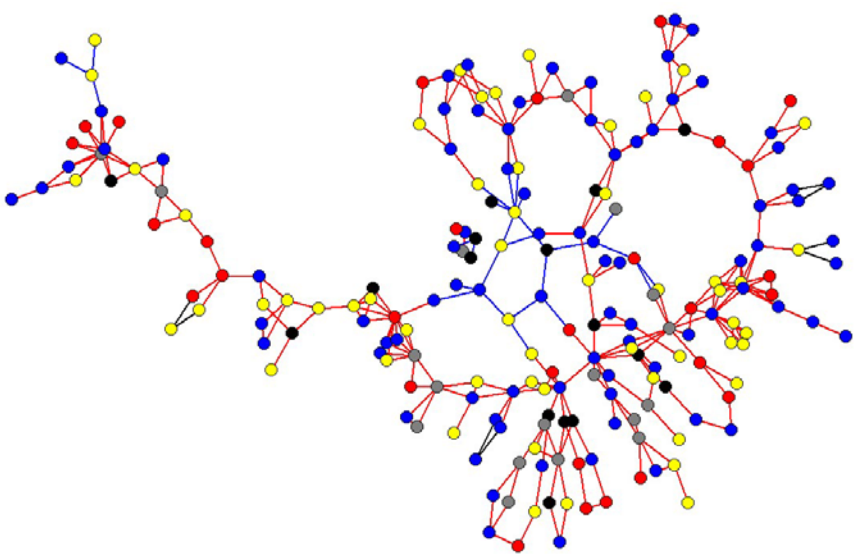


Figure 27: Zone 6 network

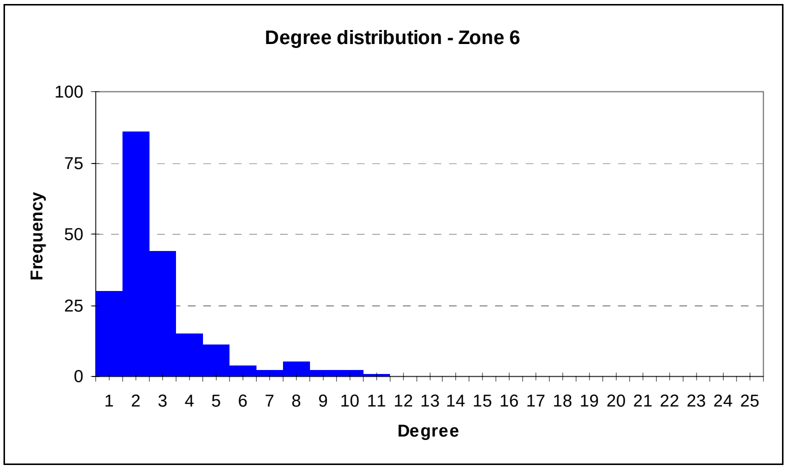


Figure 28: Degree distribution in Zone 6

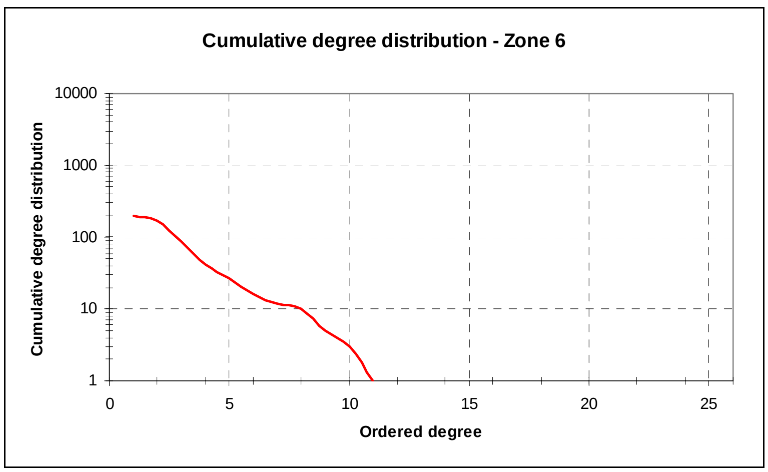


Figure 29: Cumulative degree distribution for Zone 6

	N	M	m/n	<k>	Max k	L	logn/log<k>	Alpha	C	<k>/n	R
Zone 6	137	288	2.102	2.851	11	9.500	4.696	No	0.231	0.021	-0.09

Table 14: Network characteristics for Zone 6

Zone 7

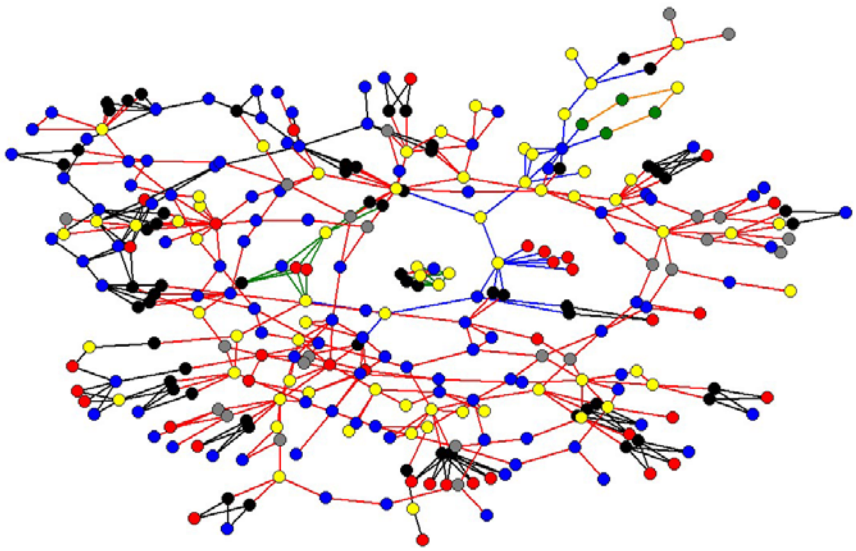


Figure 30: Zone 7 network

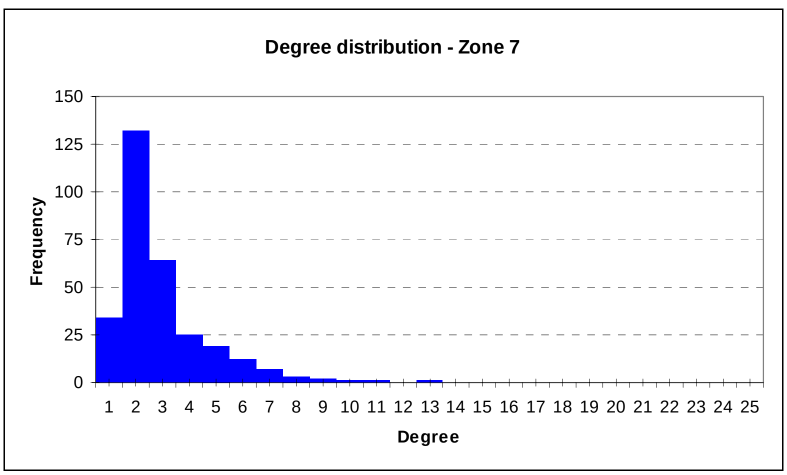


Figure 31: Degree distribution for Zone 7

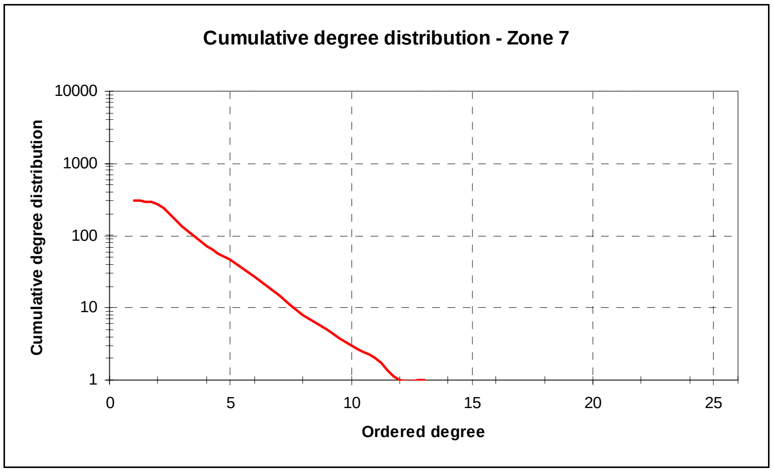


Figure 32: Cumulative degree distribution for Zone 7

	N	M	m/n	<k>	Max k	L	logn/log<k>	Alpha	C	<k>/n	R
Zone 7	301	441	1.465	2.93	13	7.902	5.309	No	0.121	0.010	-.086

Table 15: Network characteristics for Zone 7

Appendix 4: Network Analysis for Zones 4 and 8

Zone 4

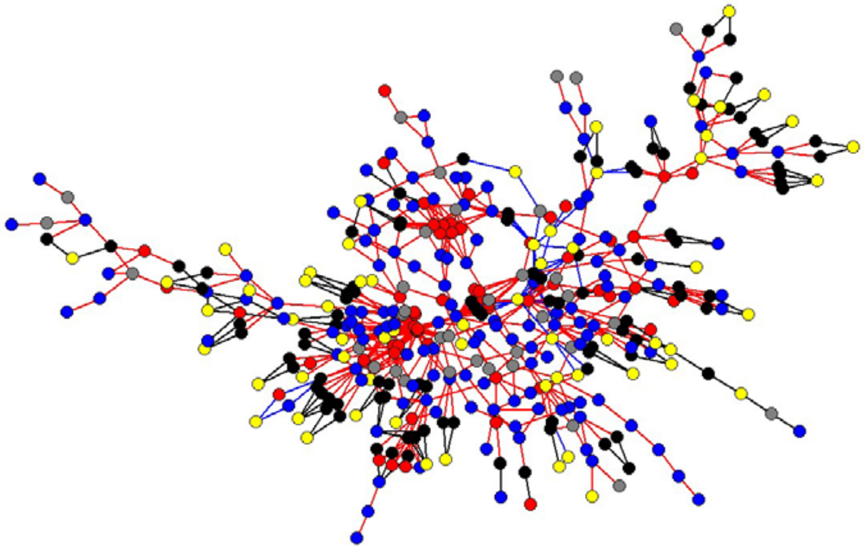


Figure 33: Zone 4 network

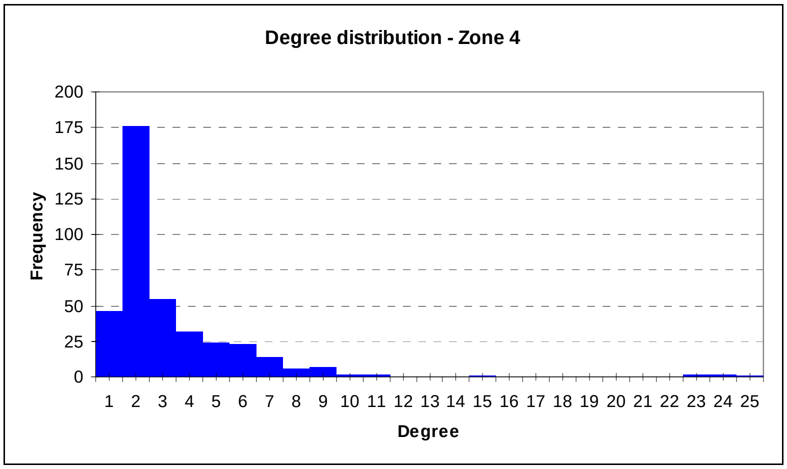


Figure 34: Degree distribution for Zone 4

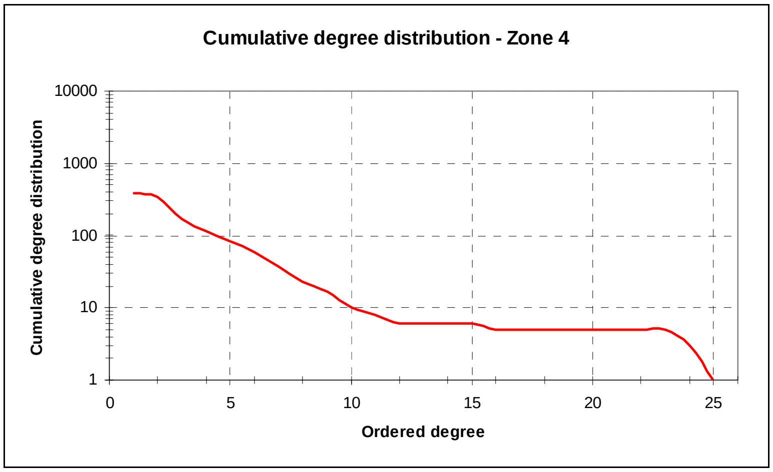


Figure 35: Cumulative degree distribution for Zone 4

	N	M	m/n	<k>	Max k	L	logn/log<k>	Alpha	C	<k>/n	R
Zone 4	393	667	1.697	3.394	25	7.853	4.889	No	0.111	0.009	0.34

Table 16: Network characteristics of Zone 4

Zone 8

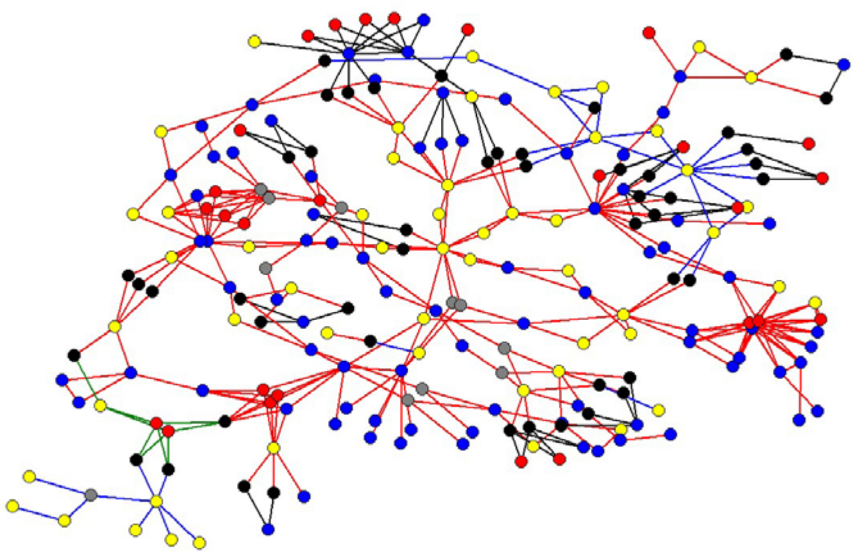


Figure 36: Zone 8 network

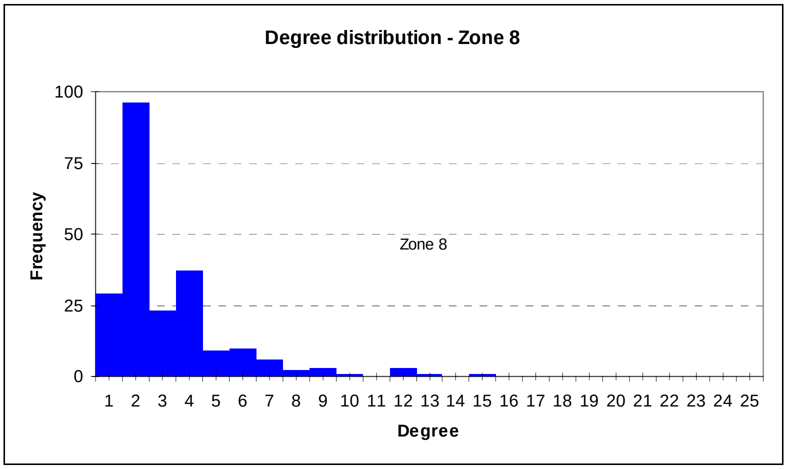


Figure 37: Degree distribution of Zone 8

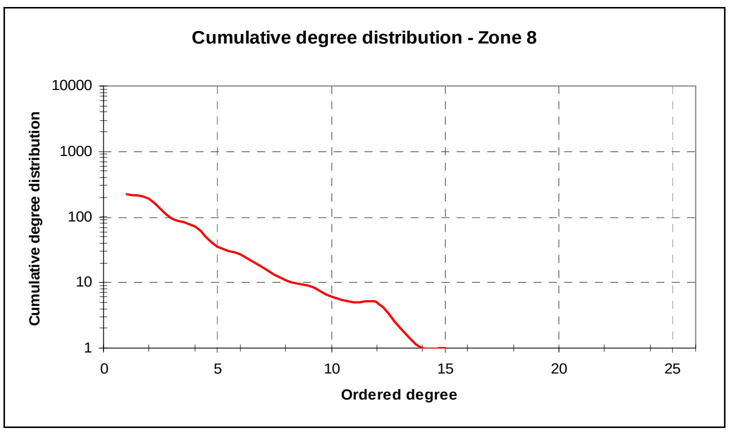


Figure 38: Cumulative degree distribution of Zone 8

	N	M	m/n	<k>	Max k	L	logn/log<k>	Alpha	C	<k>/n	R
Zone 8	221	351	1.588	3.176	15	6.282	4.671	No	0.193	0.014	0.08

Table 17: Network characteristics of Zone 8

Appendix 5: Congested zones in the overall network

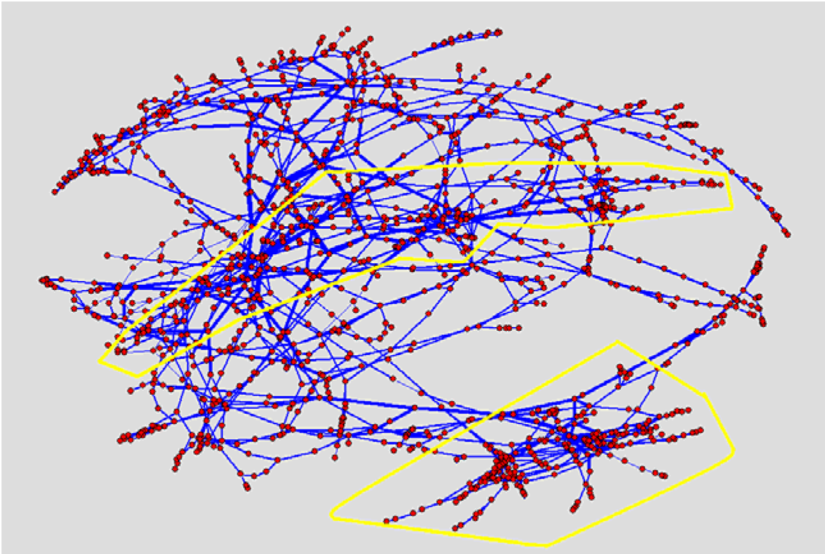


Figure 39: Congested zones in the network

Appendix 6: Effects of nodes 1100 and 1599 on the system

Through out these images we can see the cumulative effect of removing critical nodes of the system according to their highest degree first. Here it is interesting to observe that the removal will cause some smaller areas of the network to disconnect from the rest. Figure A6.1 shows the network before removing node 1100, and Figure A6.2 shows the same network after removing that particular node. We observe that the latter removal results in a big area of the system to disconnect (see blue area). The immediate effect is on the average path length, which will decrease as it is now calculated on the remaining connected areas of the system.

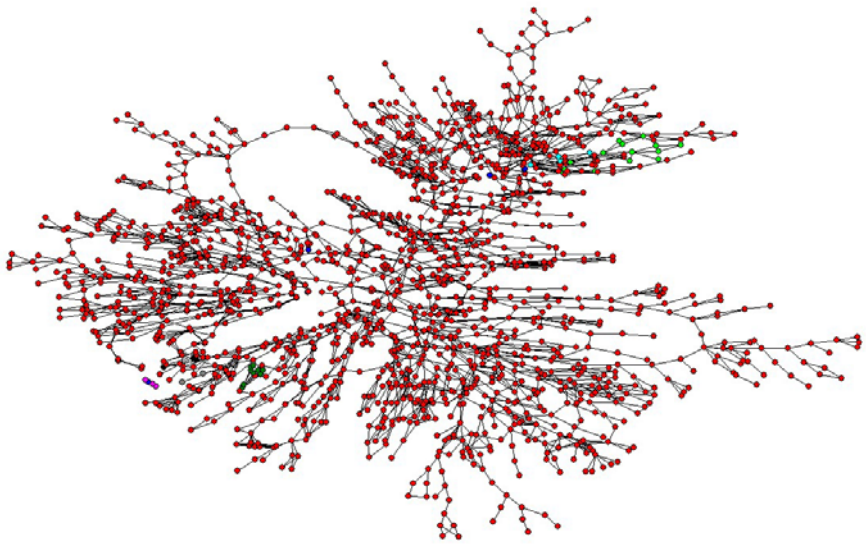


Figure 40: System representation before removing node 1100

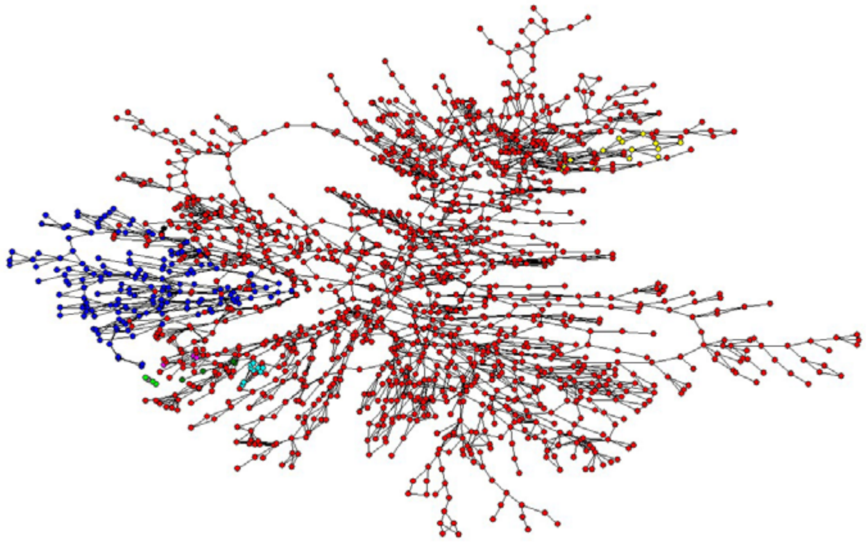


Figure 41: System representation after removing node 1100

Hare again we observe a similar effect before and after removing node 1599, Figures A6.3 and A6.4 respectively. With the latter removal, the disconnected areas get larger (see blue and black areas) and the final effect is a decrease in the average path length as it is calculated on the remaining reachable node pairs of the network.

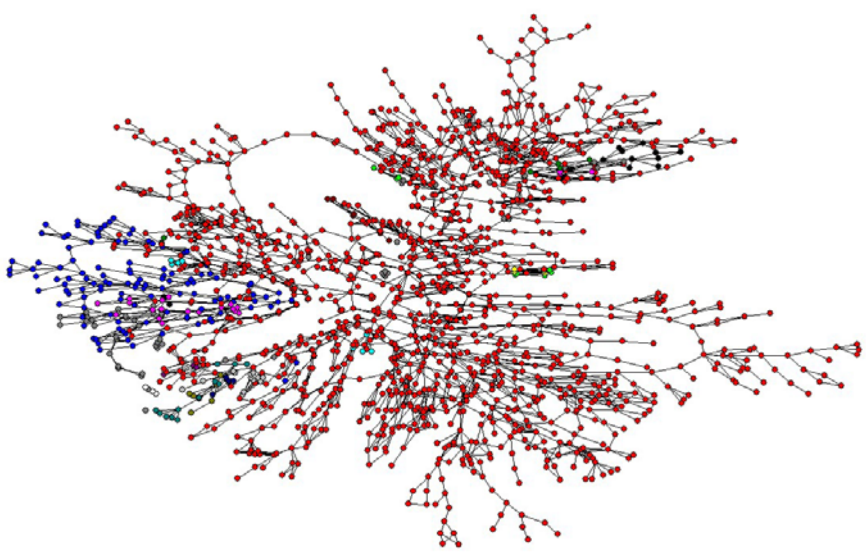


Figure 42: System representation before removing node 1599

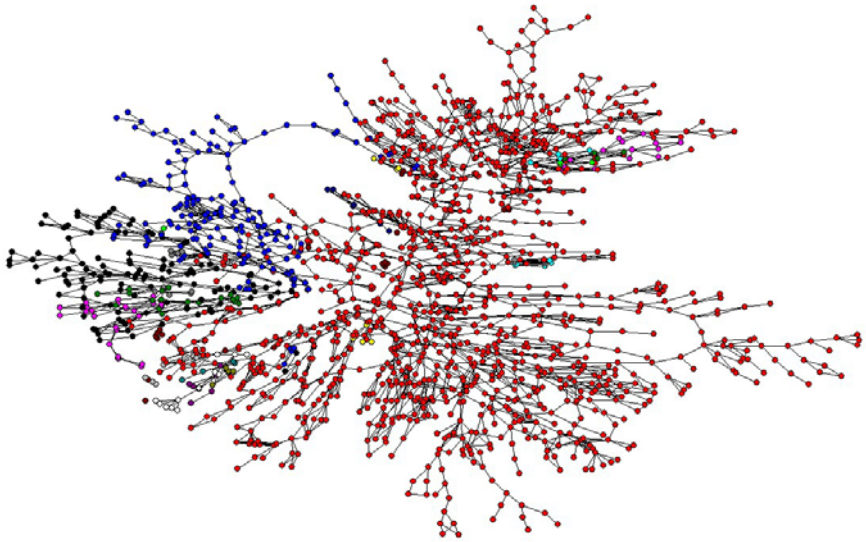


Figure 43: System representation after removing node 1599

Appendix 7: 115kV Transmission System according to Cb

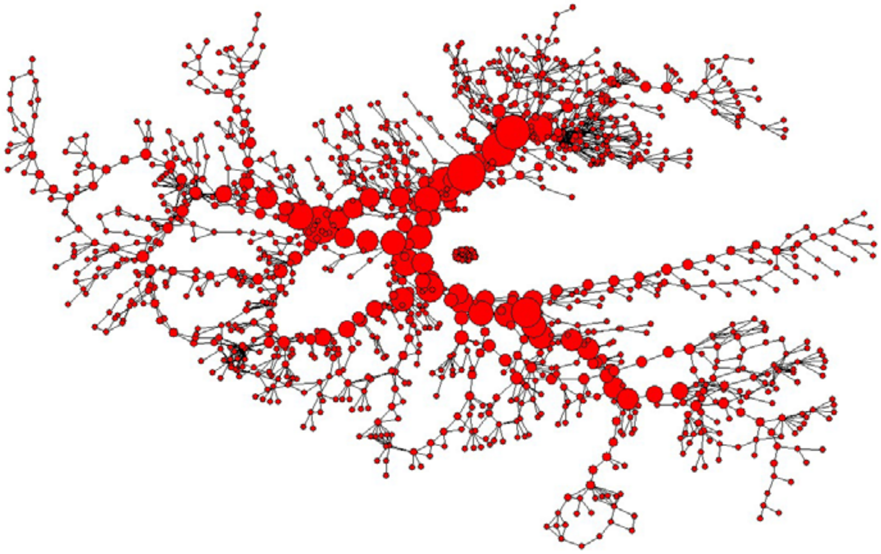


Figure 44: 115 kV system, organized by centrality

Appendix 8: Adjacency Matrix Structure

QuickTime™ and a
TIFF (LZW) decompressor
are needed to see this picture.

Figure 45: Adjacency matrix