Interdependent Response of Networked Systems to Natural Hazards and Intentional Disruptions

A Dissertation
Presented to
The Academic Faculty

by

Leonardo Augusto Dueñas-Osorio

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Civil and Environmental Engineering

Georgia Institute of Technology
December 2005

Interdependent Response of Networked Systems to Natural Hazards and Intentional Disruptions

Approved by:

Dr. Barry J. Goodno, Advisor
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Reginald DesRoches
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. James I. Craig
School of Aerospace Engineering
Georgia Institute of Technology

Dr. Ann Bostrom
School of Public Policy
Georgia Institute of Technology

Dr. Bruce R. Ellingwood
School of Civil and Environmental Engineering
Georgia Institute of Technology

Dr. Nozar G. Kishi
Catastrophe Modeling Department
AIR-Worldwide Corporation

Date Approved: November 7th, 2005

If we can understand the behavior and function of the networked systems we see around us, it will give us insight into a vast array of complex and previously poorly understood phenomena.

— Mark E. J. Newman,

The structure and function of complex networks

For my parents Rafael and Esperanza,

my sisters Sandra and Maria del Pilar,

and my girlfriend Cassy.

Instrumental in stimulating growth of my academic and social networks

AKNOWLEDGMENTS

The interdisciplinary nature of this research required the intellectual contribution of many individuals. I am indebted to the members of my doctoral guidance committee who provided continuous critical evaluation of my research activities: Dr. Barry J. Goodno, Dr. James I. Craig, Dr., Dr. Ann Bostrom, Dr. Reginald DesRoches, Dr. Bruce R. Ellingwood and Dr. Nozar G. Kishi.

I would also like to thank the fellow doctoral students from the structural engineering, mechanics and materials group of the School of Civil and Environmental Engineering. Their fruitful discussion helped me to better understand the research problems at hand.

Finally, I am forever grateful for the unconditional support of my family in Colombia. They have the patience to listen and the wisdom to offer appropriate advice.

The work reported here has been funded in part by the Mid-America Earthquake Center through National Science Foundation grant EEC-9701785.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Infrastructure systems are essential for the continuous functionality of modern global societies. Some examples of these systems include electric energy, potable water, oil and gas, telecommunications, and the internet. Different *topologies* underline the structure of these networked systems. Each topology (i.e., physical layout) conditions the way in which networks transmit and distribute their flow. Also, their ability to absorb unforeseen natural or intentional disruptions depends on complex relations between network topology and optimal flow patterns. Most of the current research on networks is focused on understanding the properties of large systems using statistical methods (i.e., the study of average network properties that emerge from the properties of its constituent elements), or on developing advanced models to capture network dynamics and evolution on small networked systems.

Despite these important research efforts, almost all studies concentrate on specific networks. This network-specific approach rules out a fundamental phenomenon that may jeopardize the performance predictions of current sophisticated models: network response is in general interdependent with other systems. The performance and functionality of a network is conditioned on the state of additional interacting networks. Although there are recent conceptual advances in network interdependencies, current studies address the problem from a high-level point of view. For instance, they discuss the problem at the macro-level of interacting industries, or utilize economic input-output models to capture

entire infrastructure interactions, or discuss the potential use of advanced monitoring and decision making units for real-time network control and stability.

This study approaches the problem from a more fundamental level. It focuses on network topology, flow patterns *within* the networks, and optimal interdependent system performance. This approach also allows for probabilistic response characterization of networked systems when subjected to internal or external disturbances. The response of real and idealized interacting systems is nonlinear with respect to the intensity of the disruptions due to the increased complexity and intractability introduced by their coupling. Methods proposed in this study can identify the role that each network element has in maintaining network connectivity and optimal flow. This information is used in the selection of effective pre-disaster mitigation and post-disaster recovery actions. Results of this research also provide guides for resilient network growth and reveal new areas for research on interdependent dynamics. Finally, the algorithmic structure of the proposed methods suggests straightforward implementation of interdependent analysis in advanced computer software applications for multi-hazard loss estimation.

CHAPTER 1

INTRODUCTION

World-wide social and economic stability is largely dependent on reliable supply of goods and services that traverse technological networked systems. Current globalization requires almost instant trading of products to equilibrate supply and demand forces (Friedman, 2005). However, this phenomenon of globalization which provides the opportunity for expanding business also has the potential to annihilate economic sectors in case of unforeseen disruptions in their productivity. Competitors are ready to provide the services or manufacturing that a particular country ceases to offer due to natural or intentional disruptions to their infrastructures.

During the last decade, the United States has experienced a series of large scale disruptions to its infrastructure revealing unexpected propagation of network disruption and unprecedented economic losses. Examples include earthquakes, satellite telecommunication failures, malicious attacks, blackouts, and hurricanes. The Northridge earthquake occurred on January 17, 1994. This moderate moment magnitude (M = 6.7) earthquake proved to be one of the costliest in the United States history with an estimate of $40 billion in direct losses alone (Eguchi et al., 1998). All critical infrastructures were severely affected delaying the recovery rate due to their interdependence. Indirect losses, with more uncertain estimates, easily doubled the direct losses figure.

The significant disruption potential of a seismic event is due to its impact felt over a large geographical region. However, there are also less severe events with far reaching consequences. In general, small disturbances are locally absorbed by the networks. But there are instances in which unfortunate sequences of events lead to the propagation of successively stronger disruptions until a significant portion of the functionality of the infrastructure system is impaired. An unforeseen satellite failure on May 19, 1998 demonstrated that malfunction of a particular element in one system can cross infrastructure boundaries due to their coupling linkages (Little, 2002). The PanAmSat Galaxy IV telecommunications satellite suffered a failure in its on-board control system. The backup switch also failed and the device rotated out of its orbital position. Over 80% of the digital pagers in the United Stated went off-line. Cable and broadcast transmissions were affected, as were credit card authorizations and Automated Teller Machine (ATM) transactions. The health care system was also critically affected as doctors and care givers could not be paged for just-in-time service delivery.

Other types of interdependencies were revealed by the terrorist attacks of September 11, 2001. In this case infrastructure destruction was geographically contained and impact on physical networks was localized. Moreover, post-disaster recovery is only efficient with substantial coordination among emergency response agencies, intelligence agencies, governmental institutions, administrative bodies and general bureaucracies (Walker, 2002; Kondrasuk, 2004). This type of interdependence, although without physical links, still operates as a network system and it is vital for optimal allocation of scarce resources.

Another example of infrastructure disruption with high-order propagation effects is power grid blackouts. Power transmission and distribution constitutes the backbone for sustained productivity. Manufacturing industries and digital-quality end users (i.e., high-technology businesses) also demand a highly reliable source of electrical current. However, the power grid can be fragile. Deviations from normal operating conditions, such as peak loads (i.e., energy demand increase), or benign disruptions from falling trees, lightning or ice storms, can result in uncontrollable cascading failures. The Western Systems Coordinating Council (WSCC) suffered a breakup on August 10, 1996. Arcs to trees progressively tripped several high voltage transmission lines near Portland, Oregon. In less than seven minutes an outage cascaded through the system, fracturing it into four islands and interrupting services to approximately 7.5 million customers on the west coast (Hauer and Dagle, 1999; Kosterev et al., 1999). Another striking blackout occurred on August 14, 2003. High power flow demands, deficiencies in management of reactive power and voltage variations, and simple lack of adherence to industry policies (e.g., inadequate tree trimming) provided the pre-conditions for the outage. The event resulted as a combination of failures in monitoring system tools, control rooms, and lack of efficient decision making—whether or not to drop 1,500 MW of load around Cleveland, Ohio to alleviate system stability. The decision was not made. As a result, 50 million people were affected in at least ten northeastern states and one Canadian province. Estimated losses reached $10 billion U.S. dollars (United States - Canada power system outage task force, 2004). Most of the physical and organizational infrastructures essential for modern societies were affected. Examples include national security; health and welfare; communications; finance; transportation; food and water supply; heating and

cooling; computers and electronics; commercial enterprises; and even entertainment and leisure.

Cyclones with the ability to impact large geographical regions are another frequent threat. Every year many tropical storms develop in the Caribbean. However, some storms evolve and gather force becoming hurricanes of category 4 or 5 strength. This unfortunate scenario became true with the category 5 hurricane Katrina, which struck Louisiana, Mississippi, and Alabama on August 29, 2005. The aftermath revealed the chaotic situations that can develop when basic societal needs are not satisfied promptly (American Society of Civil Engineers, 2005). Its initial local effects on civil infrastructure and governmental institutions resulted in the lost of about one million jobs and an estimate of damage of $100 billion. The propagation of disruptions through the interconnecting paths among social and technological networks caused an impact on fuel prices, ocean shipping and exports of agricultural products, gambling industry which is a significant source of tax revenues, and exposed the weaknesses of the government to respond to disasters (Posner, 2005; The Economist, 2005).

A clear observation common to the sample of discussed cases is that disturbances were able to efficiently spread up to the point of compromising the functionality of entire infrastructures. This propagation is in essence facilitated by the obvious underlying structure of these systems: they are tightly interconnected networks. It is clear that real networks do not operate in isolation, and that a larger problem needs to be addressed. Attempting to recognize and quantify the effects of network *interdependencies* in their response to disturbances constitutes the fundamental challenge of this research study. However, it must be recognized that these networks have also been growing

independently very fast. This growth increases their structural complexity and reduces the possibility to trace their dynamics, and the dynamics of their interconnecting systems. Fundamental approaches need to be implemented to understand the properties of these complex entities. Statistical methods for the study of average properties not of single systems, but of large ensembles of systems and their constituent parts are a promising approach to provide the basic tools for this ongoing research. This statistical approach offers a reasonable framework to discover emergent behavior and universality among systems that are composed of collections of sub-elements. That knowledge is used to understand the correlation between disturbances, network flow, topology, network response or performance, and effects of alternative improvements to prevent large scale failures.

1.1 RESEARCH GOAL AND OBJECTIVES

The concise goal of this research study is *to understand the effects that network interdependencies introduce in their response to disturbances*. The objectives to achieve the stated goal are divided in three groups:

- Network characterization
  - to investigate statistical parameters for description of network topology;
  - to identify the role that various network elements have in maintaining connectivity and optimal flow;
  - to develop performance metrics for quantifying generic network functionality;
  - to construct a power grid model for statistical analysis;

- Interdependent network performance
  - to establish governing rules for modeling network interdependencies;
  - to develop interdependent infrastructure models for statistical analysis;
  - to study the effects of systematic network disruption on independent and interdependent system performance;
  - to estimate seismic response of interdependent systems using real infrastructures of Memphis, Tennessee;

- Interdependent network improvement
  - to determine mitigation strategies that ensure continuous functionality of interdependent systems when resources are limited; and
  - to establish a research agenda for addressing additional research questions revealed by this analysis of complex system interdependencies. For instance network growth, dynamic response of networks, and system reliability.

1.2 THESIS ORGANIZATION

This thesis is divided into eight chapters. Chapter 2 presents a literature review on the recent advances in the fields of complex networks, network topology, network dynamics, and interconnected systems. Chapter 3 provides an introduction to fundamental concepts of graph theory that are used throughout the document. Basic properties to characterize network topology are reviewed, and new parameters tailored to quantify network functionality are proposed. Chapter 4 introduces the workings of

current network models, the properties of real networked systems, and the development of specific models to capture the essence of critical infrastructure topologies. Modeling of network interdependencies is presented in Chapter 5, where a novel simplified model for interconnected networks is described. Chapter 6 presents a comprehensive analysis of the response of networked systems when subjected to intentional element removal, or when subjected to seismic hazards. The effect of network interdependencies is demonstrated by comparing the response of these systems under different coupling strengths. Chapter 7 provides ideas for optimal mitigation actions, which maximize individual network stability and minimize interdependent network instabilities. Chapter 8 contains the conclusions of this research and presents relevant issues for future research.

CHAPTER 2

ADVANCES IN COMPLEX NETWORKS

Three fundamental aspects make contemporary networked systems grow in complexity. The first is the physical size and expansion of the networks to serve more end users. The second corresponds to the less tractable pattern of dependence within infrastructures to connect with other elements of their own systems. The third one corresponds to the need of the networks to interconnect to *other* systems to efficiently perform and control their operations. These network issues currently pervade all of science. However, there is one basic problem common to all disciplines: the characterization of network anatomy has not been completed, and is essential because structure *always* affects function (Strogatz, 2001).

Analysis of most complex systems indicates that they are governed by organizing principles, which should be encoded at some level in their topology and flow patterns. In the past few years there have been several developments that have promoted abundant research in the field. For instance, the computerization of data acquisition has led to the emergence of large databases that describe the topology of real networks. Increased computing power now allows investigation of large size networks and exploration of questions that were impossible to be addressed or simulated before. The breakdown of boundaries between disciplines has exposed more researchers to the notion of emerging

properties on large scale networks. And finally, there has been an increasingly voiced need to move beyond reductionist approaches and try to understand the behavior of systems as a whole (Albert and Barabási, 2002).

These trends support the search for statistical properties of large-scale networks; the development of network models that can help understand the meaning of these properties; the exploration of evolving models to find how networks become specifically structured during their growth; and the ability to predict network behavior to disturbances on the basis of measured properties and local rules governing individual elements (Newman, 2003). Recent encouraging progress has been founded on graph theory combined with statistical methods and computer simulations, and also on the insight from engineering design.

This chapter begins with a historical review of the analytical and empirical developments of graph theory. This theory has fueled, during the last decade, several advances in the calculation of network static and dynamic properties. The former properties are related to network topological structure, while the later properties are related to the response of networked systems when subjected to unforeseen disruptions, or related to their evolution and growth.

A discussion of the contributions from the field of lifeline earthquake engineering is also presented. The relevance of this field to the overall study of complex networks relies upon the extensive collection of empirical data on the performance of network elements when subjected to seismic loads. Advanced analytical predictive models have also been developed, providing unique means to validate new statistical models of

increasingly complex real systems. Loss estimation and risk analysis due to network reduced functionality are essential for today's resource allocation in disaster prevention.

This chapter also presents recent contributions from the study of a new kind of complex system: networks of *interdependent* infrastructures. A critical appraisal of the state-of-the-art in the field is also included. Readers familiar with the topic of network analysis can omit the fist two sections and focus on the last section of this chapter. That section provides the basis and justification to the issues addressed throughout the present thesis. Readers interested in an introduction to the research on networked systems can benefit from the broad discussion of the entire chapter.

2.1 HISTORICAL REVIEW

Rigorous treatment of network problems started in 1736 when Leonhard Euler offered a rigorous mathematical proof to the seven Bridges of Königsberg problem. The problem was to find a route around the city of Königsberg that would require a person to cross each of seven bridges exactly once (Barabási, 2003). He lumped the land into nodes (vertices) and replaced the bridges with links (edges), obtaining a graph with four nodes and seven links, as shown in Figure 2-1. He proved that there is no solution. A route crossing each link only once does not exist.

Further advances in the field were not made until 1959 when two Hungarian mathematicians Paul Erdös and Alfred Rényi developed the theory of *random networks*. They introduced the use of probabilistic methods to demonstrate the existence of graphs with particular properties (Erdös and Rényi, 1959). One such property is *connectivity* or the ability of the nodes to communicate with each other. The interesting aspect is that

this property and several others arise rather suddenly: while a random graph with *n* vertices and a certain number of edges is unlikely to have the property at hand, a random graph with few more edges is very likely to have the property (Bollobás, 1998a). This rapid change is called a *phase transition*. Several disciplines have used this model to understand the threshold or critical point at which the transition occurs. Researchers utilize it to understand how innovations spread, how diseases propagate, or how information flows.

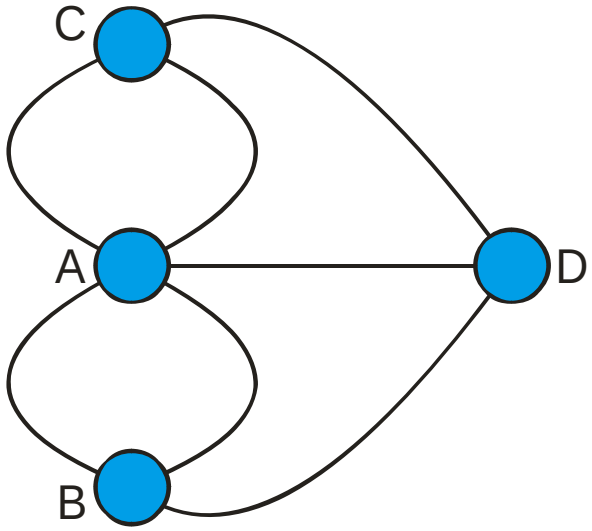


Figure 2-1. Graph representation of the seven Bridges of Königsberg problem

Information flow was the subject of a study conducted by the psychologist Stanley Milgram in 1967 on the structure of social networks (Milgram, 1967). This study revealed an important property of social networks: they are simultaneously clustered at the local scales (i.e., tight circles of acquaintances), and efficiently connected at the global scales (i.e., low degree of vertex separation). On the technological side, Paul Baran a research engineer for RAND Corporation had the task of developing

communication systems able to survive nuclear attacks. His research indicated that the ideal survivable architecture was a distributed meshlike network, redundant enough so that if some nodes went down, alternative paths maintained the connection between the rest of the nodes (Baran, 1964).

Researchers and scientists did not realized that modeling real complex networks required a shift in paradigm, despite the convenience and mathematical insights provided by random graphs models, the insights from empirical studies on social networks, and the ideas for optimal design of resilient networks. This only happened in the late 1990's, when databases from several disciplines became readily available, and general features of complex networks started to be uncovered. Sociologists, mathematicians, physicists and engineers joined forces to formally develop the new science of a connected age (Watts, 2003).

2.2 RECENT DEVELOPMENTS

The new science of networks is a multi-disciplinary field where fundamental theoretical advances are required to model real phenomena. The laws of physics, engineered laboratory experiments, and pre/post disaster network evaluations provide the tools to validate and calibrate models. This body of knowledge enables asking questions about more complex problems, such as that of propagation of disruptive effects through the intricate structure of interdependent systems. The rigorous study of networked systems relies upon the principles of graph theory and the convenience of statistical methods. Applied engineering provides the boundary conditions, initial conditions, and operational regimens for successful implementation of theoretical models; and the

multiple disciplines interested in the topic constitute the beneficiaries and the source of feedback for reshaping theoretical developments and applications.

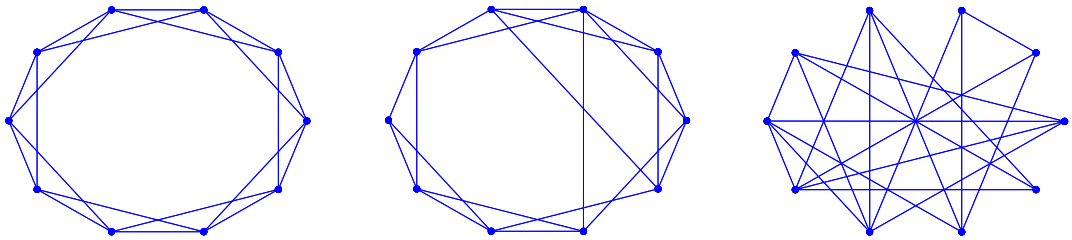
2.2.1 Graph Theory and Statistical Methods

With the urge to address realistic problems, researchers turned to observational evidence to support their model developments. In few years the field has moved from measuring topological properties for classification of networks, to successfully developing models that reproduce real network properties. The field is also constructing dynamically evolving models whose elements age, grow, and interact with each other. Research in this new area is shedding some light on the mechanisms that enable networks to efficiently operate and respond to large unexpected demands.

2.2.1.1 Static Structure of Networks

The first network model capable of simultaneously displaying high local clustering and a low degree of global separation (as social networks have) is the *small-world* network (Watts and Strogatz, 1998). Construction of this model suggested the mechanism that enabled real social networks to display those properties: just a few elements of clustered groups at the local scale need to be linked with other distant tight clusters. The algorithm to develop small-world graphs simply starts with a regular network or *lattice* where every node has the same number of neighbors. Then, for a graph with n vertices, each link is *rewired* at random with probability p. If $p = 1$ the resulting network is a random graph because every link is rewired (Watts, 1999). Figure

2-2 shows the graph with $n = 10$ vertices, four neighbors per vertex and three values of the rewiring probability p which increases randomness as $p \rightarrow 1$.



(a) (b) (c)

Figure 2-2. Small-world graphs with (a) $p = 0$, (b) $p = 0.1$, and (c) $p = 1$

An additional property of the small-world networks is that every vertex tends to have, on average, the same number of links. However, empirical data have shown that in most known networks (e.g., internet, world wide web, research collaborations, protein regulatory networks, cellular metabolism, business, etc), the number of nodes versus the number of links per node follows—denoted as vertex degree, $d(v)$— follows a power-law in the tail of their distributions: $p_{d(v)} \sim d(v)^{-\beta}$ for some constant exponent β. These tails imply that a few nodes are abnormally well connected while most other nodes only have a few connections. Some geographically distributed technological networks such as power grids and transportation networks are the exception since they show exponential tails in their vertex degree distributions (i.e., most nodes have similar number of connections). The first model that reproduced the power-law distribution of the vertex degree was referred to as the *scale-free* network (Albert et al., 2000; Barabási and Bonabeau, 2003). The mechanisms that allow the model to display such a power law

distribution are network growth and preferential attachment. Rather than starting with a set of *n* connected vertices, each disconnected vertex is joined to the network with links connecting to the vertices that are already connected. This connection is done with a probability proportional to the number of links that each existing connected vertex has. Figure 2-3 shows a sample scale-free network with $n = 10$ vertices, grown by connecting each new node to two existing nodes following a preferential rule.

Additional research has shown that it is possible to develop networks with power-law distributions without using preferential attachment (Kaiser and Hilgetag, 2004). The mechanism in this case is growth of the network in a given space, and letting some nodes to be pioneers and be positioned far from the original group. This generates multiple interconnected clusters which a few nodes accumulate a large number of links.

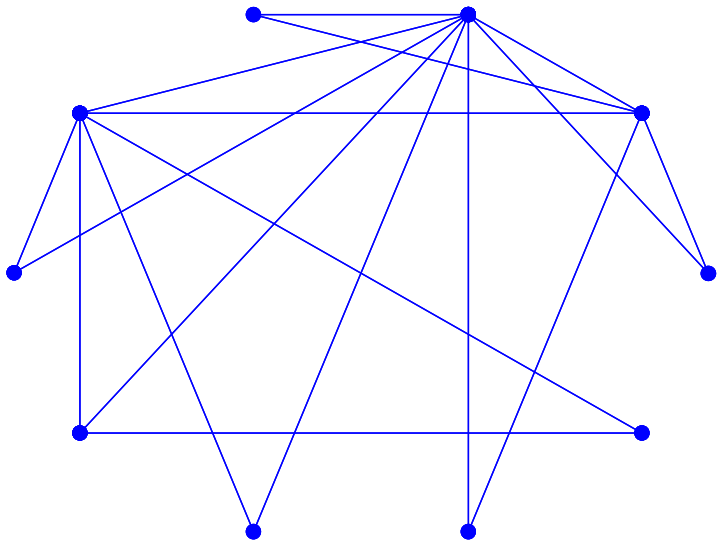


Figure 2-3. Topology of a scale-free network

The fact that simple rules helped explain the topology of complex networks initiated a research effort to find properties and the mechanisms that produced them. The aim has been to determine *universal* laws that govern the development and functionality of complex natural and technological networks. An example uncovering universality is the study of scientific collaboration networks (Newman, 2001a). Scientists are nodes, and links among them are established if they have co-authored one or more papers. Researchers used databases of scientific papers in physics, biomedical, and computer science. They found that from all authors in a discipline 80% to 90% of them had co-authored a paper. This is referred to as a *component* or cluster of connected nodes. Also, the scientific collaboration networks displayed a large local clustering. This supported the hypothesis that social networks form community cliques due to their common interests. Other statistics that are useful to the understanding of how the networks operate are typical distances between scientists and measures of *centrality* (i.e., relative importance of network elements according to their role in network topology and flow.) This network shows a remarkable short distance between any two scientists—representing a true small-world effect—which grows logarithmically with the number of authors, n, in the network (Newman, 2001b). These networks were also used to compute a parameter called *betweenness*, which helps identify who are the most influential people in the network or the ones who control the flow of information between most others. Additional centrality measures suited for specific flow processes have appeared in more recent studies of generic networks used for transfer of goods, money, e-mail, and infections (Borgatti, 2005; Newman, 2005). These new approaches account for the kinds of trajectories taken by the flow (e.g., shortest paths, trails, or walks), and the method of

spread (e.g., transfer, replication, or broadcast). Trails and walks are just different ways to traverse a network not necessarily taking the shortest routes. Network *efficiency* is one parameter that relies upon measurement of shortest paths between all vertices (Latora and Marchiori, 2001).

Generalizations of known graph models have led to the development of random networks with arbitrary *vertex degree* distribution (i.e., the number of links per vertex). Traditional random graphs, as proposed by Erdös and Rényi, possess a Poisson vertex degree distribution. Exact expressions for generalized random graphs have been derived for the position of the phase transition (i.e., when a random network becomes suddenly connected by adding links); for the size of the largest cluster after the transition; and for the vertex-vertex distance within a graph (Newman et al., 2001). Classical statistical methods are also used to study families of network models derived by requiring the expected properties of a graph ensemble to match a given set of measurements of real-world networks (Park and Newman, 2004). This approach is mathematically and conceptually sophisticated, and has the potential to provide predictive models that explain several features of real networks, rather than having several models each explaining a particular phenomenon.

More conventional advances are incremental improvements of existing models. For instance, traditional scale-free networks exhibit power law-degree distribution for the number of connections per node, short vertex to vertex distance, and low local clustering. Generalizations have been made to create scale-free networks with tunable local clustering which allows generation of networks with all the properties of the scale-free

network class, but in agreement with the observed high local clustering of real networks (Holme and Kim, 2002).

Characterization of real networks has benefited from these theoretical advances. One example is the network of web pages and hyperlinks connecting them—the world wide web (WWW). Studies indicate that the *page rank*, or measure of authority of the WWW to prioritize pages matching a query, concentrates on a small number of nodes. Network analysis says that this is a consequence of the average *in-degree* (i.e., number of links entering into a node, if the network is directed) of the network (Nakamura, 2003). Self-links and multiple links may be the responsible for high page rank of non-relevant web pages. Another example of real network characterization shows that transportation networks possess small-world characteristics. Specific examples include subway systems (Latora and Marchiori, 2002) and railroad networks (Sen et al., 2003). Universal properties have also been proposed for general distribution networks (e.g., drainage basin of a river, cellular metabolism, or civil transportation infrastructure). Research results reveal that the size and the flow rates of efficient transportation networks have a linear relation in log-log scales (i.e., power-law relationship) irrespective of dynamical or geometric assumptions (Banavar et al., 1999). Other researchers use empirical data from real networks to refine predictive capabilities of existing network models. In the case of the internet topology, scale-free networks reproduced several of the observed properties, such as vertex degree and average node separation. However, in order to enhance their predictive capabilities to more characterization parameters, scholars have modified the preferential attachment rule from linear on the number of links (i.e., proportional to the vertex degree) to nonlinear relations on the number of links (Zhou and Mondragón, 2004).

The availability of refined models to reproduce real network properties, and the sophistication of the mathematical tools to formulate them allow addressing different questions. These questions go beyond network description and focus on prediction. However, predictive capabilities require understanding of how the shape of the various systems conditions the way flow traverses them.

2.2.1.2 Dynamic Processes on Networks

Knowledge of network topological properties and the mechanisms that produce them has facilitated the study of *network dynamics* (i.e., processes going on within networks). Intuitively, researchers have claimed that network topology affects the way in which any flow traverses a network. Today there are tools for their rigorous treatment. The concept of *percolation*, borrowed from physics, describes the process in which vertices or links on a network are randomly designated "occupied" or "unoccupied." Figure 2-4 illustrates percolation of nodes and links, referred to as *site* percolation and *bond* percolation, respectively. Site percolation indicates whether or not network nodes have failed (i.e., occupied state). Bond percolation indicates the state of network links. This idea has been extended to address fundamental dynamic processes such as epidemics within social networks, or cascading failures in technological networks—where infection or failure is the "occupied" state. Approximate expressions for the percolation probability (e.g., susceptibility to infection) have been proposed for disease propagation on small-world networks (Newman and Watts, 1999). Percolation processes on random graphs with arbitrary vertex degree distribution have been suggested as models for the robustness of communication or distribution networks to breakdown or

sabotage. Researchers have found exact solutions to this particular problem, providing another tool to make predictions about the behavior of networked systems under general types of breakdowns (Callaway et al., 2000).

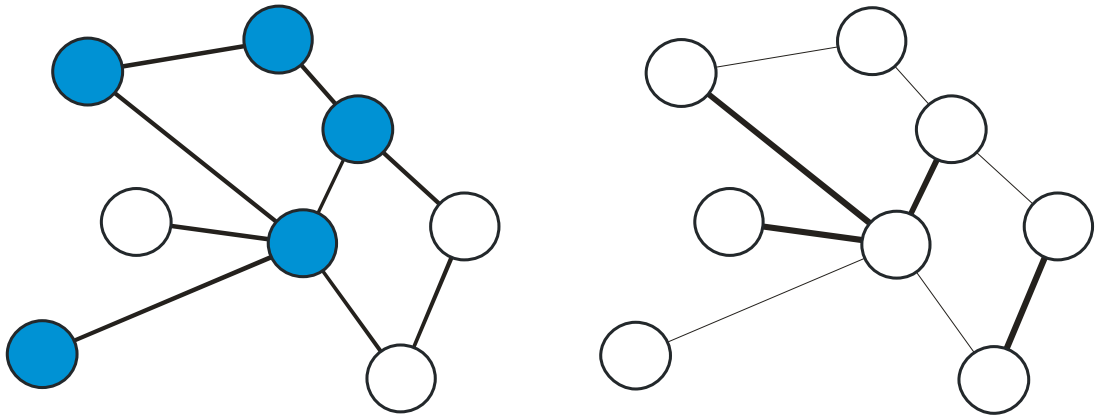


Figure 2-4. Site and bond percolation on a network

Percolation has also opened the door to development of simple, yet powerful dynamic models to predict cascades on networks. Using arbitrary random graphs, investigations have demonstrated that simple threshold rules govern the decisions of interacting nodes based on the state of their neighbors (Watts, 2002). Node connectivity and node stability are the parameters that most influence the susceptibility to cascading failure. Dynamic processes taking place on scale-free networks have demonstrated that the breakdown of a single node is sufficient to collapse the functionality of entire systems if the node is among the ones with the largest *load*—the amount of flow traversing it (Crucitti et al., 2004). Theoretical estimates supported with empirical evidence have also been obtained for the phase transition point—onset of total collapse—in terms of a

characterizing parameter of node capacity (Zhao et al., 2004). Data transport in scale-free networks has also led to important observations. It has been found that a power law distribution governs the amount of data packets passing through a vertex when every pair of vertices sends and receives data following shortest paths (Goh et al., 2001).

Additional studies have addressed the dynamics of networks taking place on various graph topologies such as branching trees or regular lattices in one and two dimensions (i.e., rings and meshes, respectively). Models for communication networks reveal that when nodes deliver packets of information proportionally with increments on the load traversing them, the network is not susceptible to collapse due to congestion of information. When the number of delivered packets of information decreases with load increments, there is a sudden phase transition from a functional state to a collapsed state; and when the number of delivered packets of information is independent of the load traversing them, the phase transition is a continuous function from functionality to collapse (Guimera et al., 2002). Theoretical work indicates that there is a critical traffic load above which the probability of traffic congestions destroying the network communications capabilities is finite (Moreno et al., 2003). When the traffic flow is proportional to the number of shortest paths passing through the nodes, the congestion conditions (i.e., critical points below which the traffic flows free and above which traffic congestion occurs) are independent of network size and topology (Zhao et al., 2005). Similar studies have been performed to measure the size of *avalanches* or cascades in simplified power transmission systems where solution of circuit equations is not included in the models. Threshold rules alone are used to activate protective devices in each node. It has been shown that the probability distribution of the size of the disturbance has

power law scaling (Sachtjen et al., 2000). Slight disruptions to elements that have a critical role in load delivery can cause overload failures. Heterogeneous distribution of loads in networks—as in the case of the internet or the power grid—is a precondition to large-scale cascades because heavily loaded nodes operate close to their capacity, and if overloaded, the amount of flow to be redistributed easily overwhelms their neighbors (Motter and Lai, 2002).

After the increase in awareness to possible attacks to critical infrastructures, research has also focused on uncovering what would happen to network functionality if particular network elements are intentionally removed. Answers to this type of question are being pursued in the context of network models and real networks. The common starting point is to determine the importance of the network elements, so that the top ranked ones are removed first. Importance can be defined in terms of the role an element has to maintain connectivity or to transmit flow. After every removal the status of the network in term of some functionality parameter is monitored. Studies for random graphs, small-world graphs, and scale-free graphs with and without high local clustering have revealed important characteristics. One such result is that random graphs and small-world graphs are equally vulnerable to random or to targeted attacks, whereas scale-free graphs are very resilient to random disturbances, but also highly vulnerable to targeted attacks (Holme et al., 2002). Their vertex degree distribution plays an important role in determining network vulnerability. Disruptions to links also show that scale-free networks are more sensitive to attacks on *short-range* links—where range is the length of the shortest path between two nodes in the absence of that link (Motter et al., 2002).

Network functionality parameters have been evolving from measuring the size of the remaining largest cluster after every network element removal, to quantifying the ease of connection between any two nodes of the network. A more refined parameter, referred to as network efficiency, measures the ease of node to node connection after disruptions but taking into account the Euclidean distance between them (Crucitti et al., 2003). Other network performance measures have been developed to quantify the functionality of classes of networks, such as the ones for flow transmission from generation nodes to distribution nodes. *Connectivity loss*, a parameter to measure the ability of the distribution nodes to remain connected to generation nodes, has been introduced to study the effects of disturbances on power grids (Albert et al., 2004).

Research has also focused on development of strategies for mitigation of network vulnerability and enhancement of network recovery. Biological networks show that the nature of the recovery after a crash (e.g., time to recovery), depends upon the organizational structure that survives the crash (Jain and Krishna, 2002). This is also true for technological systems and for network models. Some researchers have proposed models of networks for complex system design that display the connection effectiveness of scale-free networks and the robustness to random attacks of exponential networks (Shargel et al., 2003).

Practical implementation of mitigation actions requires exploration of optimal strategies in terms of low cost and high impact. This examination implies exhaustive analysis of alternatives and determination of the protection policy that maximizes the improvement of network performance given a damage that minimizes network performance (Latora and Marchiori, 2005). This approach marks one of the state-of-the-

art methods for spotting critical network elements and for outlining network expansion plans.

Observed network damage after major disasters has provided some evidence of recurring failure modes and particular element vulnerabilities. Therefore, it has been necessary to bridge the gap between the various elements representing the abstract set of nodes and links in graph theory and their real physical and mechanical properties. This has attracted the interest of engineers, and the general approach consists of quantifying and describing the *fragility* of network elements (i.e., susceptibility to fail given certain level of disruption), estimating their importance for infrastructure functionality, and measuring the potential effects of mitigation actions on system reliability.

2.2.2 Applied Engineering

Lifelines represent the set of civil infrastructure systems essential for continuous operation of modern societies. These systems are distributed in large geographical extensions, are composed of nodes and links, and are suitable for analysis using network and optimization theory. However, due to the significant uncertainty in the parameters that influence the response of network elements, engineers have adopted probabilistic methods for performance estimation. Early attempts included the development of probability matrices, based on expert opinion, to estimate damage to facilities or likelihood of breaks per unit length in linear facilities such as pipelines or high voltage transmission lines (Applied Technology Council, 1991). The focus of their research was on the response to seismic hazards. This approach considered the network elements to be *independent*. Refinements to this methodology have been the research motivation for

numerous predictive response models, experimental studies, and probabilistic methodologies for risk quantification.

2.2.2.1 Analytical, Numerical and Empirical Developments

Recent studies have focused their research at the individual and system level elements. These studies are validated and calibrated with empirical evidence collected by post-disaster reconnaissance teams and experimental tests. An example of individual element studies includes dynamic and static analysis of structures for transmission of electric energy (Riley et al., 2003). This study identifies failure modes for lattice steel towers and for steel poles—tension and compression capacity of legs, and base moment capacity, respectively. Other studies include interaction among various pieces of closely integrated equipment, for instance, the voltage transformer–bushing system. In this system the bushing is a porcelain or fiberglass insulating element that connects transformers to the external power lines. Finite element models are developed to estimate the dynamic response of the system, and provide evidence that the expected response of transformers—more flexible than originally thought—can exceed acceptable limits and response predictions from design guidelines (Institute of Electrical and Electronics Engineers, 1997; Matt and Filiatrault, 2003). Another example is the investigation of dynamic interaction between connected electrical substation equipment with dissimilar characteristics. Hysteretic models for enhanced rigid connectors (i.e., bus sliders) are proposed to improve prediction of dynamic response of interacting equipment items (Song et al., 2003). Simplified models to estimate interconnection point displacements have also been developed (Dastous and Filiatrault, 2003). Finally, interacting power

transmission towers with high voltage transmission lines prove to have a critical separating span beyond which their interaction is highly detrimental due to the considerable mass that lines add to the systems (Li et al., 2003).

Models to capture the entire flow dynamics of a power grid have also been in development. Transient stability programs—step by step integration of circuit differential equations—are the most reliable. There are other approaches that rely upon approximations and are part of expert systems, pattern recognition methods, or neural networks. These methods are useful for the analysis of large scale problems and steady state responses. Transient stability approaches are more suitable for dynamic response evolution. For instance, they are used to reproduce power outages (Kosterev et al., 1999). Software development with advanced visualization tools, such as WinIGS-F, also performs power flow analysis of electric power systems. Specifically, it allows the user to model any three phase power system together with its ground wires, and analyzes the performance of the system under steady state conditions (Advanced Grounding Concepts, 2005). Similar tools have been developed for other transmission and distribution systems. For instance, the steady-state analysis of potable water flow using the principles of conservation of mass—supply and demand equilibrium—and conservation of energy—head losses through the system must balance at every point (Hatestad, 2004).

Empirical estimates of network damage have also been widely used, especially in pipeline analyses. Damage ratios in repair per kilometer for common pipeline materials, pipe diameters, and pipe support conditions became the preferred fragility measure for linear infrastructures (O'Rourke and Ayala, 1993). Experimental studies using shake tables and pseudo-dynamic testing have also been undertaken to expand the

knowledgebase of network element response (e.g., steel storage tanks, pumps, electrical substation equipment, compressors, etc.) Some projects have already revealed the inadequateness of current electrical equipment evaluation procedures (Schiff and Kempner Jr., 2003).

Observed damage has also influenced definition of performance objectives for lifelines. An example for water systems includes guidelines for acceptable performance requirements in networks exposed to diverse natural hazards (Eguchi and Honegger, 2003; Graf et al., 2003). More ambitious efforts try to unify the design of key utility and transportation systems (e.g., electric power, telecommunication, water, waste water, oil, natural gas, rail, and shipping ports) to achieve a desired performance in natural hazards. A fast evolving set of guidelines has been that for seismic design and retrofit of piping systems (American Lifelines Alliance, 2002). The ultimate goal has been to guide lifeline system operators on how to ensure intended network *functionality* after earthquakes (e.g., position retention, leak tightness, operability, etc.) This functionality demanded additional research beyond prediction of the response of network elements. It required mapping element response with potential service impact on end-users.

2.2.2.2 Loss Estimation and Mitigation Methods

Network element response aggregation is essential for estimation of system-level performance after disruptions. The uncertainty in the response and the resistance of network elements is significant and needs to be included in loss estimation methods. Reliability analysis of network systems provides a framework to assess within certain confidence bounds the probability of failure of network elements (Yang and Shaoping,

2003). Some studies, in addition to accounting for the reliability of the elements, also consider flow capacity to enhance reliability assessment. A particular example for water delivery systems recommends the construction of reliability surfaces (i.e., reliability contour maps) to guide maintenance decision-making (Quimpo and Wu, 1997). In the case of highway systems, research has suggested simple performance parameters such as length of operating network, or accessibility between points to estimate the reliability of complete systems (Chang and Nojima, 2001; Lleras-Echeverri and Sánchez-Silva, 2001). Additional analytical efforts are invested in development of efficient analytical algorithms for generic large scale system reliability evaluation (Li and He, 2002).

Other probabilistic risk assessment methods use Monte Carlo simulations to produce failure sets at an element level, and then monitor network performance (Ostrom, 2003). This methodology has been widely accepted because for moderate order networks—number of vertices ≤ 10,000—it is feasible to perform exhaustive evaluation of element failure. Other researchers, focusing on seismic hazards, have included liquefaction potential and have used Geographic Information System (GIS) technology to estimate network performance (Hwang et al., 1998). GIS also provides a useful link between network service reduction from estimated damage and the differential social impact that reflects demographic diversity—age, income, ethnicity, etc. Two example applications correspond to the water delivery and transportation networks of Shelby County, Tennessee (French and Jia, 1997; Werner et al., 1997).

Visualization can enhance understanding of the extent of damage after a loss estimation exercise. Important efforts to aid city planners, emergency agencies, and decision-makers concentrate in development of software applications for estimation of

potential losses from disasters. A module for damage estimation of lifelines and utility systems is available in the Federal Emergency Management Agency's tool for loss estimation, HAZUS-MH (Federal Emergency Management Agency, 2005). This program allows for deterministic and probabilistic damage calculation at different levels of accuracy. Confidence levels depend on the quality of the input data and the chosen estimation model (e.g., statistical aggregation, simplified network flow analysis, or detailed network flow analysis). Another software project with a network analysis module is the Mid-America Earthquake Center's tool for risk management, MAEViz (Mid-America Earthquake Center, 2004). This software, still under development, combines risk communication, advanced visualization, and dynamic decision-making. The goal is not to be simply a loss estimation tool but rather a system to engage stakeholders to manage their risks and help them minimize the consequences of disasters by implementing effective mitigation measures. Software development for specialized networks is another driving research force. A particular tool for water systems permits benefit-cost analysis, comparison of mitigation strategies, and prioritizing capital for improvement programs (Huyck et al., 2003). For power networks, seismic risk management systems are also under development (Shumuta, 2003). This approach allows for network mitigation benefit-cost analyses including financial costs of business disruptions. Benefit-cost exercises that include treatment of costs that vary over time enhance the predictive capability of the analyses (Chang and Seligson, 2003).

Almost all loss estimation and consequence minimization methodologies evaluate the damage when the systems reach equilibrium after the disruption. However, some research teams are looking at the residual capacity of the systems to estimate the duration

of the service interruption (Nojima and Sugito, 2003). The study of network *resilience*—the ability of network elements to recover from disturbances—constitutes one of the major research areas in lifelines engineering. Post-earthquake restoration processes are traditionally characterized by restoration curves (i.e., recovery of functionality as a function of time). However, recovery processes are intimately dependent on the recovery rates of interacting systems (Cagnan and Davidson, 2003). This notorious effect has contributed to the already significant need for exploring interdependencies among networked systems.

2.2.3 Network Interdependencies

Large scale natural disasters and contemporary terror threats have triggered a wake up call for understanding the complexities of interacting critical infrastructures. Experts advising the government of the Untied States of America produced a comprehensive report highlighting the need for mathematical unifying frameworks that allow modeling and prediction of the response of networked systems (National Research Council, 2002). This report includes an explicit request for communicating across disciplines. This requires domain experts to learn one another's languages in order to pose significant questions and interpret the answers usefully. This request also implies supporting systems-level thinking in education to create a generation of scientists who understand the interconnectedness of modern society's constituent parts. The federal government has also initiated a coordinated effort to disseminate and implement a national strategy for securing the infrastructures and assets vital to security, governance, public health, economy and public confidence (The White House, 2003). This strategy

delineates specific responsibilities and provides guiding principles to both government institutions and private sectors. It asks for setting goals to appreciate and harness the complexities of interdependent mixes of facilities, systems and functions.

From the engineering point of view, interdependencies among civil infrastructure systems are growing in number and strength. Some infrastructures are becoming markedly more essential than others. Electric energy and information transfer (e.g., telecommunications) are leading the stability, control, automation, efficiency, and reduced operation costs of most infrastructure services (Heller, 2001). Emerging frameworks to model interacting infrastructures are just starting to be developed, and they are already facing multi-disciplinary integration challenges.

One of the simplest approaches to model interdependencies is sequentially ordering the implications of infrastructure failure. Stages include identification of the causes of perturbation; definition of the possible infrastructure failure; identification of the potential disruptive events derived from infrastructure failure, and quantification of the damage. Each stage of the sequence is connected to the preceding and following stages by a probabilistic function based on the frequency of occurrence for any two linked stages (Little, 2002). This method has the drawback of having a linear structure that differs from the parallel interaction of real infrastructures.

A better approach is one that accounts for simultaneous interaction of collections of network elements in which change occurs with time. Rules for interaction among evolving network elements depend on several dimensions. For each system, there are differences in the types of failure, the strength of the coupling with other networks, the environment in which they operate, the type of interdependency, the state of operation,

and the characteristics of the infrastructure (Rinaldi et al., 2001). *Complex adaptive systems* (*CAS*) have been proposed as a viable framework to deal with interacting infrastructures (Axelrod and Cohen, 1999). In a CAS there are many kinds of participants or *agents* (e.g., network elements), who interact in complicated ways that continuously reshape the future. Each agent is an entity with location, capabilities and memory (Gell-Man, 1994). These properties aggregated at a system-level generate behavior and responses that go beyond the simple addition of agent contributions. This *emergent* behavior is the hallmark of a CAS. Interesting conceptual depictions of CAS technology envision their use not only for predictive purposes, but also for self-healing of infrastructure systems (Amin, 2000; Amin, 2001). It is claimed that if agents sense any anomalies in their surroundings they can work together, essentially reconfiguring the system to keep the problem local. Additional CAS applications use idealized networks (e.g., small-world) to represent infrastructures whose interdependencies are governed by empirical rules. Such an approach allows combination of physical infrastructure networks with organizational networks, creating a stronger model to predict societal impacts (Comfort et al., 2004). In essence, this method captures the interconnection between technical systems that provide public services and the organizational systems that manage them. However, incomplete information, complex data mining of available information and high subjectivity in trend interpretations and governing rules for evolution have prevented CAS approaches to grow faster in practical applications.

Other frameworks, with the objective of providing reasonable global estimates of the effects of interdependencies, use ideas from input-output models of the economy which are inherently interdependent. An example of these studies focuses on the

estimation of economic losses to a specific industry sector caused by damage to three other infrastructure systems (i.e., power, water, and gas), which are evaluated separately. Then, using an input-output (I-O) impact methodology, interdependencies are accounted for in the analysis of exchanges of goods and services among industries (Chang et al., 1996). Another example is an extension of Leontief's economic model which enables understanding and forecasting of the effects on one segment of the economy upon changes in another (Haimes and Jiang, 2001). In this model, entire infrastructure systems take the place of economic sectors. It is a high-level approach sensitive to the definition of the amount of inoperability induced by one infrastructure system on another system. Advances to the same model have been recently proposed. New features allow for dynamic interactions and different temporal recovery patterns (Haimes et al., 2005a; Haimes et al., 2005b). However, the interconnectedness is still modeled among infrastructure *sectors* without dealing with elements of each infrastructure. Recent attempts to combine economics—which provides bounds to the interactions—and elements of an infrastructure network are founded on *game* theory. In this framework, coupled networks are treated as layers of transportation networks (e.g., vehicles, freight, or data) comprised of non-cooperative game players—Cournot-Nash dynamic agents. The goal is to solve the problem for flow equilibrium and optimal budget allocation (Zhang et al., 2005). This approach, despite requiring complex behavioral rules of interaction, uses a holistic perspective that can lead to a more efficient understanding and modeling of interconnected infrastructure systems.

Even though there have been significant conceptual and theoretical advances in the field of interdependent networks, most frameworks use highly simplified real

networks, or theoretical network models to formulate the interdependencies problem. These frameworks do not explicitly ask for the role that the layout of the chosen network has in flow distribution. They do not account for the simultaneous *topology-flow* effects on response prediction. Network *topology* affects the way in which any *flow* traverses it, and vice versa. Few studies utilize a well defined topological layout to let network dynamics occur within them. One recent example uses two dimensional idealized grids as layouts for network interactions (Newman et al., 2005). However, further improvements can be done to better represent the natural topology of infrastructure networks, and more importantly, to include the configuration of network interdependent elements in their layout.

2.3 CRITICAL REVIEW OF CURRENT RESEARCH

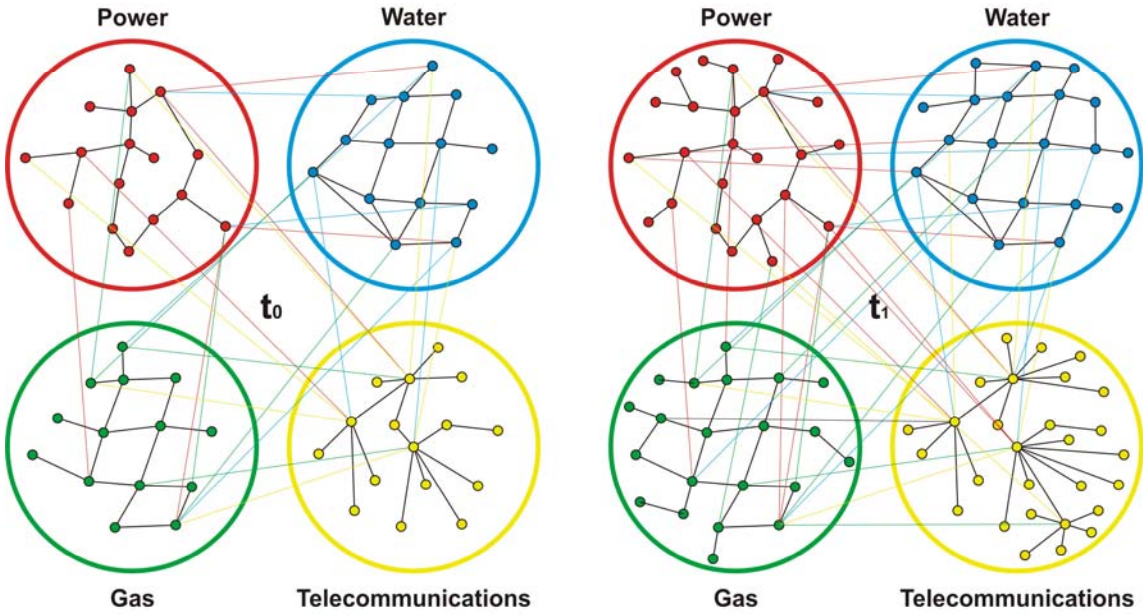
The body of knowledge on the new science of networks is strengthening its fundamental basis and this is allowing an expansion of applications. Statistical network characterization is providing important evidence about the universality of network class properties. Also, when these properties are reproduced by analytical or algorithmic models, researchers are able to obtain significant insights on the mechanisms that produce them. Engineering is providing relevant experimental information regarding the fragility of individual and interacting network elements. Sophistication in specific network system models is increasing network response prediction. Loss estimation methods now allow for dynamic decision-making and realistic representation of the social impacts induced by network malfunctions. In the infrastructure interaction area, there are significant advances to allow moving from conceptual recognition of the

problem to the development of models that account for these coupling effects. Interacting economic sectors provide evidence of the relevance of accounting for infrastructure coupling, and complex adaptive systems (CAS) show the potential to become the future tool for dynamic representation of evolving interacting networks.

Despite these advances in less than a decade, statistical characterization of infrastructure networks still needs parameters to quantify network performance and the impact on society if malfunctions occur. Parameters to measure network redundancy or to quantify their service flow reduction need to be investigated. Refinement of the probabilities of failure of network elements given different hazards needs to be included in the calculation of network response to disturbances. Explicit accounting of the effects of network topology should also be considered in the development of engineering reliability assessment and mitigation plans.

The topology and flow pattern characteristics of interdependent systems also require more investigation. This will provide the understanding of the role that each network element has in maintaining individual and interdependent network connectivity and flow. Further analysis of this type of information will lead to optimal strategies to improve the performance of modern multi-infrastructure systems. Initial steps to address this needed research have been conceptually proposed by the author (Dueñas-Osorio et al., 2004). This preliminary study suggests an evolving model for encapsulating the topology and flow characteristics of critical infrastructures such as power, water, and gas that share a common geographical area. Figure 2-5 sketches an expanded version of this evolving interdependent network. The figure illustrates that individual network growth occurs from the core of the existing network to the peripheries—proportional to

population growth, and that the interdependencies increase uniformly throughout the entire space.



(a) at time t_0 (b) at time t_1

Figure 2-5. Interdependent networks in the same spatial domain at times $t_0 < t_1$

Additional ongoing research is expanding the understanding of the effects that network disturbances, network topology and optimal engineered service flow patterns have on interdependent response (Dueñas-Osorio et al., 2005c; Dueñas-Osorio et al., 2005b). Specific tasks of this research examine the static properties of networks, identify consistent performance measures useful for characterizing generic network functionality, and investigate approaches for efficient mitigation actions that account for network flow type, topology, spatial interconnection, strength of coupling, and direction of the interconnectedness.

CHAPTER 3

NETWORK PROPERTIES

Every networked system exhibits a distinct *topology* or physical layout. Depending on the nature of the network, its structural configuration usually evolves to optimize the type of flow to be transported. In the case of civil infrastructure, transportation networks exhibit a mesh-like structure (Figure 3-1a); telecommunication networks display a decentralized topology (Figure 3-1b); the World Wide Web (WWW) shows an disproportionate case of decentralization in which a few highly connected vertices are responsible for network stability (Figure 3-1c); and the power grid at the transmission level displays a sparse mesh-like structure—due to its geographical extent—while at the distribution level it exhibits a radial tree-like structure (Figure 3-1d). Other utility networks such as potable water and natural gas—if gas has widespread consumption—have a less sparse mesh-like structure within urban centers, and a very sparse (e.g., minimally connected) topology for interurban transmission. Oil networks also have a minimally connected tree structure.

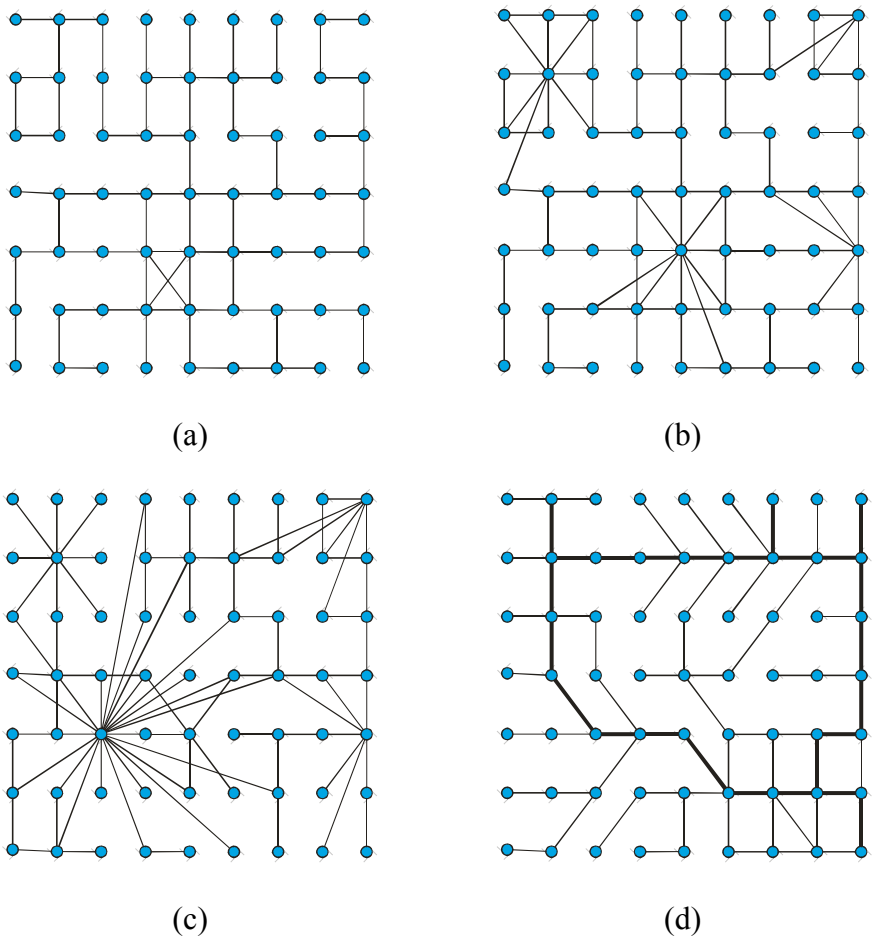


Figure 3-1. Sample of network topology diversity: (a) transportation, (b) telecommunications, (c) WWW, and (d) electric power

In addition to topological differences, networks display distinctive features in terms of their mechanisms for flow dynamics. In most civil infrastructure networks the flow process has a fixed destination from a particular origin, and it tends to happen along the most efficient route—the *shortest path*. This need for optimal flow traversal sets the basis for determining and proposing measurable network properties that simultaneously account for their physical topology and their flow processes along the routes of least resistance. Examples of flow processes different from transmission of goods starting at node *i* and ending at node *j* are the kinds of traffic exhibited by broadcasting and e-mail.

These processes allow movement of data *packets* from node to node, but at some point certain nodes make flow diffusion by replication, as opposed to flow diffusion by transfer. They take their incoming packet and replicate it for massive distribution. When social networks are analyzed, a larger variety of flow processes is observed (Borgatti, 2005; Newman, 2005). Money traverses a social network via *random walks* (i.e., a bill can move in any direction from person to person, and it is allowed to revisit nodes and links). A particular private rumor, unlike dollar bills, can simultaneously flow in several places. However, one person does not tell the same person the same story twice, even though a person can hear the same rumor from several other people. This makes the flow a *trail* rather than a walk. Infections flow like gossip, from person to person, but do not re-infect anyone who already has had it—and survived—because they became immune. If immunization did not take place, the person can be re-infected, making the flow a *trek*—which is a walk that does not backtrack.

Concepts from modern *graph theory* are fundamental to enable measuring of these observable differences in network topology and flow types. Their unambiguous characterization is essential for classification and establishment of correlations between performance, network structure, and flow patterns. This chapter presents a discussion of the concepts of graph theory that are necessary to build up tools for network characterization. It also introduces strategies to identify the importance of network elements according to their role in connectivity and flow. Additionally, it presents network performance measures which provide different levels of detail to estimate network functionality and impact on end users after internal or external disruptions. This

chapter also emphasizes the development of algorithms that efficiently calculate network properties and performance measures in computer applications.

3.1 INTRODUCTION TO GRAPH THEORY

Standard graph theory terminology used throughout the following chapters is introduced in this section. A set $A' = \{A_1,\ldots,A_k\}$ of disjoint subsets of a set A is a *partition* of A if $A = \bigcup_{i=1}^{k} A_i$. A set of all *k-element* subsets of A is denoted by $[A]^k$. A *graph* or *network* is a pair $G = (V,E)$ of sets satisfying $E \subseteq [V]^2$; thus, the elements of E are 2-element subsets of V (Figure 3-2). The elements of V are the *vertices* (i.e., nodes) of the graph G, whereas the elements of E are its *edges* (i.e., links). The number of vertices of a graph G is its *order* denoted by $|G|$ or n, and the number of edges is its *size* denoted by $\|G\|$ or m (Bollobás, 1998b; Diestel, 2000).

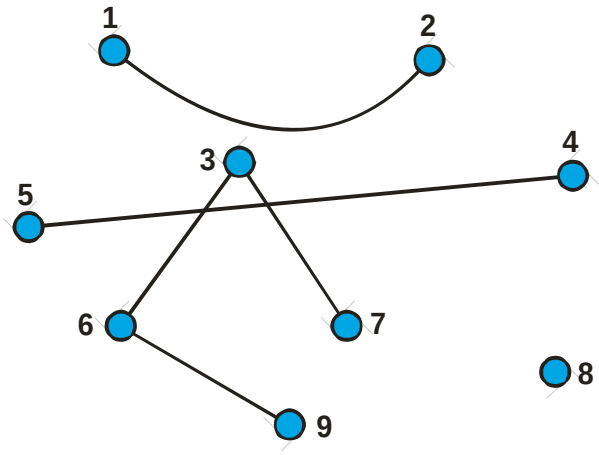
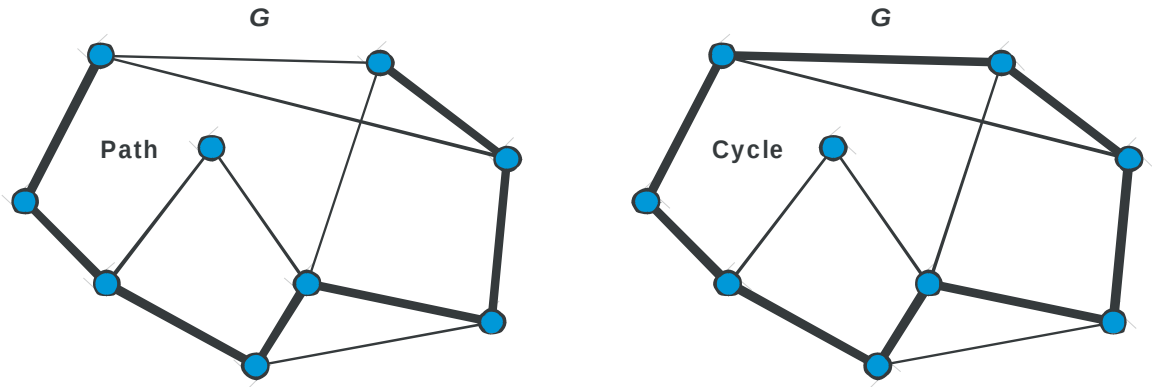


Figure 3-2. A graph on $V = \{1, \ldots, 9\}$ with edge set $E = \{ [1,2], [3,6], [3,7], [4,5], [6,9] \}$

A vertex v is *incident* with an edge e if $v \in e$; then e is an edge at v. The set of all the edges in E at a vertex v is denoted by $E(v)$. Two vertices x, y of G are *adjacent*, or *neighbors*, if xy is an edge of G. If all the vertices of G are pair-wise adjacent, then G is *complete*. A complete graph on n vertices is a K^n; for example a K^3 is a triangle. If $V' \subseteq V$ and $E' \subseteq E$, then G' is a *subgraph* of G, written as $G' \subseteq G$. A subgraph $G' \subseteq G$ is a *spanning* subgraph of G if V' spans all of G, i.e., if $V' = V$.

A *path* is a non-empty graph $P = (V,E)$ of the from $V = \{ x_0, x_1,\ldots, x_k \}$, $E = \{ x_0x_1, x_1x_2, \ldots, x_{k-1}x_k \}$ where the x_i are all distinct. The vertices x_0 and x_k are *linked* by P and are called its *ends*. The number of edges of a path is its *length*, and a path of length k is denoted by P^k. In paths, k is allowed to be zero, thus $P^0 = K^1$. If $P = x_0\, x_1 \ldots x_{k-1}$ is a path and $k \geq 3$, then the graph $C := P + x_{k-1}x_0$ is called a *cycle*. The length of a cycle is its number of edges or vertices and is denoted by C^k (Figure 3-3).



(a) (b)

Figure 3-3. Paths and cycles in bold: (a) a path $P = P^7$ in G, and (b) a cycle C^8 in G

The *distance* $d_G(x,y)$ in G of two vertices x, y is the length of a shortest x—y path in G; if no such path exists, $d_G(x,y) := \infty$. The greatest distance between any two vertices

in *G* is the *diameter* of *G* denoted by diam(*G*). A vertex is *central* in *G* if its greatest distance from any other vertex is as small as possible. This distance is the *radius* of *G*, denoted by rad(*G*). These distance measures are topological parameters that do not include network performance information such as the rate of flow transmission.

A non-empty graph *G* is called *connected* if any two of its vertices are linked by a path in *G*. A maximal connected subgraph of $G = (V,E)$ is called a *component* of *G* (Figure 3-4). A component, being connected, is always non-empty.

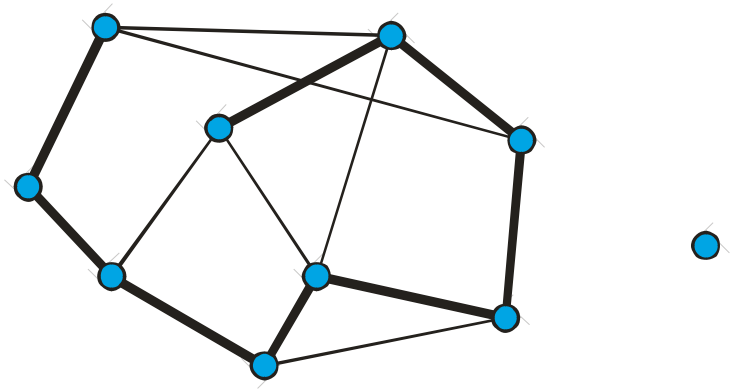


Figure 3-4. A graph with two components and minimal spanning connected subgraphs

An *acyclic* graph is a graph that does not contain any cycles, and it is called a *forest*. A connected forest is called a *tree*—a forest is thus a graph whose components are trees. The vertices of a tree that have only 1 incident edge are its *leaves* (Figure 3-5).

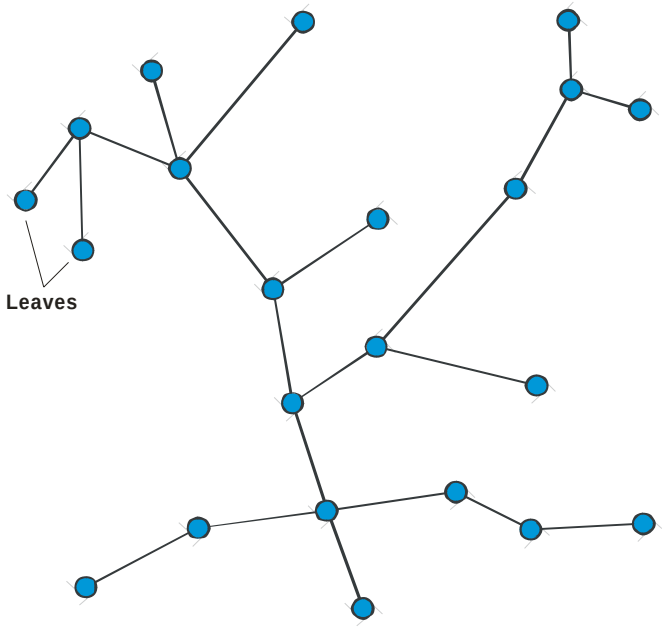


Figure 3-5. A tree as a minimally connected graph

The *incidence matrix* $B = (b_{ij})_{n \times m}$ of a graph $G = (V,E)$ with $V = \{ v_1, \ldots, v_n \}$ and $E = \{ e_1, \ldots, e_m \}$ is defined by:

$$b_{ij} := \begin{cases} 1 & \text{if } v_i \in e_j \\ 0 & \text{otherwise} \end{cases} \qquad \textbf{(3-1)}$$

Finally, the *adjacency matrix* $A = (a_{ij})_{n \times n}$ of G, which constitutes its more fundamental mathematical representation, is defined by:

$$a_{ij} := \begin{cases} 1 & \text{if } v_i v_j \in E \\ 0 & \text{otherwise} \end{cases} \qquad \textbf{(3-2)}$$

3.2 FUNDAMENTAL PROPERTIES

Several parameters can be calculated to characterize the *topology* (i.e., geometric configuration) of a network. This section presents some of the most relevant parameters currently used in the field, and introduces a proposed parameter that will be used to estimate the potential resilience of a network.

3.2.1 Mean distance, *L*

The network *mean distance*, *L*, has become one of the fundamental parameters in complex network analysis. Its importance is due to its ability to measure whether or not a network has the *small-world effect*—that is, most pairs of vertices are connected by a short path through the network. *L* for undirected graphs is defined as:

$$L = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \geq j} d(i, j) \qquad \textbf{(3-3)}$$

where $d(i,j)$ is the shortest distance from vertex i to vertex j, and n the order of the graph. If the number of vertices within a distance k of a typical central vertex grows exponentially with k, then the value of L increases as log n. Networks are said to show the small-world effect if the value of L scales logarithmically or slower with the network order, n, for a fixed average number of edges per vertex.

The definition of L is problematic in networks that have more than one component. In such cases, there exist vertex pairs that have no connecting path, and hence d is assigned with an infinite distance, which drives the value of L to also be

infinite. An alternative is to define *L'* as the reciprocal of the *harmonic mean* of *G*—the average of the reciprocals (Newman, 2003):

$$L' = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \geq j} \frac{1}{d(i,j)} \qquad \textbf{(3-4)}$$

where the infinite values of $d(i,j)$ contribute nothing to the sum. In essence, these parameters measure the effectiveness of the network elements to communicate at a global scale on the order of *n*.

3.2.2 Vertex degree, *d*(*v*)

The *degree* $d_G(v) = d(v)$ of a vertex *v* is the number $|E(v)|$ of edges at *v*. This is equal to the number of neighbors of *v*. A vertex of degree 0 is *isolated*. The number $\delta(G) := \min \{ d(v) \mid v \in V \}$ is the *minimum degree* of *G*, while the number $\Delta(G) := \max \{ d(v) \mid v \in V \}$ is the *maximum degree*. If all vertices of *G* have the same degree *k*, then *G* is *k-regular*. The *average degree* of *G* is simply:

$$d(G) = \frac{1}{|V|} \sum_{v \in V} d(v) \qquad \textbf{(3-5)}$$

Analysis of the vertex degree for theoretical and real graphs has revealed that the distribution of their vertex degree plays a crucial role in determining the fate of a network when subjected to random or targeted attacks (Albert et al., 2000). If the vertex degree

distribution is Possion with short thin tails, then the network is equally vulnerable to random or malicious disruptions because all vertices have a typical degree. If the vertex degree distribution follows a power law with long thick tails (i.e., distribution in which most vertices have a typical degree, but that a few vertices have a disproportionately large degree), then the networks show a significant resilience to random disruptions. However, they also exhibit a dramatically vulnerable response if the disturbance is directed to the vertices with the highest degree.

3.2.3 Clustering coefficient, γ

In order to define the clustering coefficient of a graph, it is necessary to first introduce the concept of *neighborhood*. The neighborhood, $\Gamma(v)$, of a vertex v is the subgraph that consists of vertices adjacent to v without including v itself (Figure 3-6). The neighborhood $\Gamma(S)$ of a connected subgraph S is the subgraph that consists of all vertices adjacent to any of the vertices in S, but not including the vertices of S. In the special case that $S = \Gamma(v)$, $\Gamma(S) = \Gamma(\Gamma(v)) = \Gamma^2(v)$. Then, the clustering coefficient, γ_v uses $\Gamma(v)$ to characterize the extent to which vertices adjacent to any vertex v are adjacent to each other (Watts, 1999). More precisely:

$$\gamma_v = \frac{\left|E(\Gamma_v)\right|}{\frac{1}{2}d(v)(d(v)-1)} \qquad \textbf{(3-6)}$$

where $|E(\Gamma_v)|$ is the number of edges in the neighborhood of v and the denominator represents the expansion of a binomial coefficient which amounts to the total number of

possible edges in $\Gamma(v)$. This clustering coefficient can be regarded as a local measure of connectivity—it measures how connected the network is in local scales (i.e., in the order of $|V(\Gamma_v)| << n$).

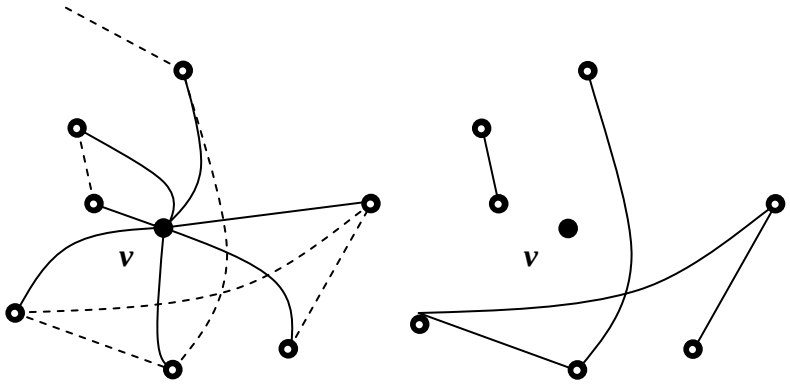


Figure 3-6. The neighborhood $\Gamma(v)$ of a vertex v: (a) neighbors of v, and (b) links among neighbors of v

3.2.4 Redundancy Ratio, R_R

Large scale complex networks, such as the power grid, may be sporadically subjected to infrequent natural disasters (e.g., earthquake, hurricanes, tornadoes, etc.), or to malicious attacks. However, these networks are also constantly subjected to direct disturbances of their elements, either by frequent natural hazards (e.g., windstorms, ice, fire, lightning, etc.), vandalism, or simple malfunction due to aging or improper operating conditions. Most of the frequent disturbances are locally absorbed by the networks, and the end-users remain unaware of their occurrence (Institute of Electrical and Electronics Engineers, 2004). This fact results from the ability of the networks to redistribute the flow at the location of the disturbance. In other words their global stability depends on the

capacity to manage and redistribute local disturbances. When this does not work, the various infrastructure networks can head to imminent cascading failures, traffic congestions, or widespread pressure losses.

This local disturbance absorption motivated the author to propose a parameter that captures the redundancy of the network at local levels. Then, if G is a graph, the *redundancy ratio* of a vertex, R_{Rv}, is a parameter that counts the number of independent paths from a vertex $v \in V(G)$, to each of the vertices of the set of the neighbors of its neighbors, $V(\Gamma^2(v))$. This count is normalized by the maximum possible number of node-independent paths from v to the vertices of $V_S(\Gamma^2(v))$, where the graph $S = \{ v \cup \Gamma(v) \cup \Gamma^2(v) \}$ is a *complete* graph. Let $I(i,j)$ denote the maximum number of node-independent paths (i.e., paths that only share vertices i and j) between each pair of *distinct* vertices (i,j) in G. This counting can be efficiently computed using approximating algorithms (White and Newman, 2001). Figure 3-7 shows a graph G with $I(i,j) = \min\{d(i), d(j)\}$, which is true for *any* graph. The redundancy ratio of a vertex v in a graph G is (Dueñas-Osorio et al., 2005c):

$$R_{Rv} = \frac{1}{(|S|-1)^2} \sum_{j \in V(\Gamma_v^2)} I(v, j) \qquad \textbf{(3-7)}$$

where $(|S|-1)^2$ is the number of independent paths between a vertex v and the vertices of the vertex set $V_S(\Gamma^2(v))$ in a complete graph on $|S|$ vertices. The most redundant simple graph is a complete graph. Hence, the median of R_{Rv} for all $v \in V(G)$ results in a measure of the redundancy ratio for the graph G, where $0 \leq R_R \leq 1$, and values close to 1 identify highly redundant networks. $R_R = 0$ indicates completely fragmented graph (i.e., no edges left),

which implies that the number of *components* equals the order n. $R_R = 1$ indicates a complete graph. For large n, paths have $R_R \to {}^1/_8$, and stars (i.e., graphs with n-1 leaves and 1 vertex connecting them all) have $R_R \to 0$.

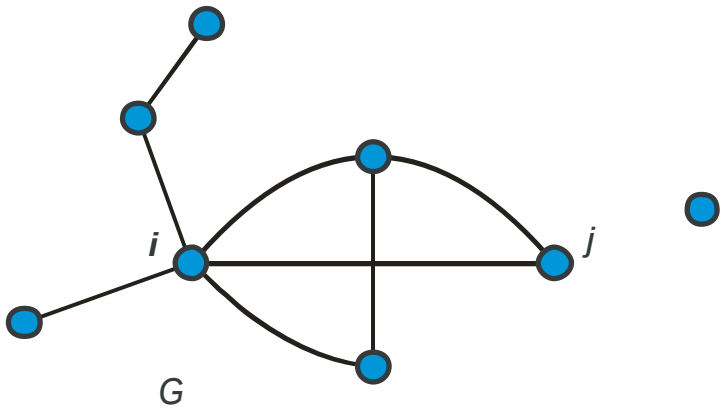


Figure 3-7. Graph with $I(i,j) = \min \{ d(i), d(j) \} = 2$

If a large scale disturbance occurs (one that is not locally contained) and the number of disturbed vertices is comparable to the order n of the network G, then it can be expected that graphs with high R_R will still display a higher global redundancy. This correlation comes from the fact that in most graphs the number of vertices reachable from a vertex v increases exponentially with the distance from it. In other words, high R_R implies that the number of vertices in the i^{th} neighborhood of v for $i > 1$ grows exponentially with i, and therefore the likelihood of path existence between vertices far apart is also high.

Measuring network properties, despite being theoretically feasible, still poses challenging implementation restrictions. Large complex systems are composed of thousands of elements that computationally restrict the number and type of

characterization parameters. Only efficient algorithms—which imply simple topological parameters—are useful for computer applications, since otherwise, the amount of time to run them becomes unbounded.

When the characterization of the properties of any network can be established by running algorithms within a computer application, it is fair to think about the feasibility of parameter calculation when the networked systems are large in order and size. Appendix A introduces the concept of *computational complexity* to address the performance of computer algorithms as a function of network growth. This appendix is aimed to illustrate the feasibility of implementation of network analysis in software applications (e.g., tools for loss estimation and risk management of multiple hazards on networked systems). The appendix includes a summary table with the worst-case time required by an algorithm to solve a particular problem. The running time is generally expressed as a function of n and m and is represented by the "big O" notation (Table A-1).

The source code for the major network properties and performance measures is included in Appendix B.

3.3 NETWORK ELEMENT RANK-ORDERING

The fundamental elements of an infrastructure network are its vertices and its edges. Both are crucial for shaping the structure and influencing the properties exhibited by a network. However, what would happen to the network in the absence of a particular element? To address this question it is necessary first to determine the set of elements that would be more likely to be absent or removed by malfunction, natural hazards or attacks. This section focuses on defining the criteria to establish the importance of

network elements based upon their role on connectivity, centrality, and flow transfer. The discussion is restricted to rank-ordering the importance of network *vertices*. Each vertex removal impairs the functionality of one or more edges, whereas each edge removal impairs the functionality of exactly one edge. In other words, vertex removal is more effective at being harmful to network performance.

Three parameters are introduced to rank order the importance of a particular vertex $v \in V(G)$ for the functionality of a network G. This rank ordering is relevant because networks tend to display avalanches or cascading failures after disruption of high ranked vertices. The propagation mechanism can compromise all elements, in descending order, of the rank-ordered sets. In general, the cascading phenomenon occurs when a particular element of a network initially ceases to provide its intended function—due to either internal or external disturbances. Then, following managerial strategies seeking maximal returns and minimal operation costs, the *flow* traversing that node is redistributed to adjacent nodes which usually function close to their maximum capacity. Some of those adjacent nodes may be operating so close to their capacity that the additional inflow from the redistribution can induce a failure. This process repeats itself until either the network absorbs the disruption locally, or, if the set of initial conditions are all unfavorable, until a large scale disruption propagates further, thus compromising a significant portion of the network and its functionality.

Rank ordering strategies are presented in increasing order of information demand and significance of their ordering criteria.

3.3.1 Vertex degree rank-ordering

This strategy is based upon the *vertex degree*, $d(v)$, of a vertex $v \in V(G)$. This parameter $d(v)$, computed for all v, permits determination of its frequency distribution. Most networks in nature such as social, informational and biological, display power law vertex degree distributions where a few highly connected nodes coexist with many minimally connected nodes (Barabási, 2003; Barabási and Bonabeau, 2003). However, some of the man-made technological networks exhibit Poisson vertex degree distributions. A relevant example is the power grid of the western United States (Watts and Strogatz, 1998; Watts, 1999).

Regardless of the shape of the histogram of vertex degrees, it is safe to say that the most connected vertex, represented by the vertex v with maximum $d(v)$ or with $d(v) = \Delta(G)$, plays an important role in keeping the network as a *component*. Rank ordering by vertex degree provides an ordered set for investigating what would happen to the networks if their most connected vertices are removed one at a time. This strategy is purely topological and requires only the information contained in the adjacency matrix A_G of a graph G. The amount of time required to establish $d(v)$ for all $v \in V(G)$ is $O(n)$. Selecting the largest value from the set takes time $O(1)$. If the set is recalculated after every removal (i.e., another n times), which is reasonable since the topology of the graph changes when a vertex is removed, then the worst-case complexity is $O(n^2)$.

3.3.2 Vertex betweenness rank-ordering

The *betweenness,* B_v, of a vertex v is defined as the total number of shortest paths that pass through vertex v, when the shortest paths are calculated between every pair of

vertices $(i, j) \in V(G)$, and v is not considered an *end* of any shortest path (Newman, 2001b; Borgatti, 2005). An illustrative extreme example is to consider a graph which can be disconnected by removing just one edge or *bridge* (Figure 3-8). The vertex betweenness of vertices x and y would be amongst the highest of all $V(G)$ because any flow sent from one side to the other has to pass through x and y.

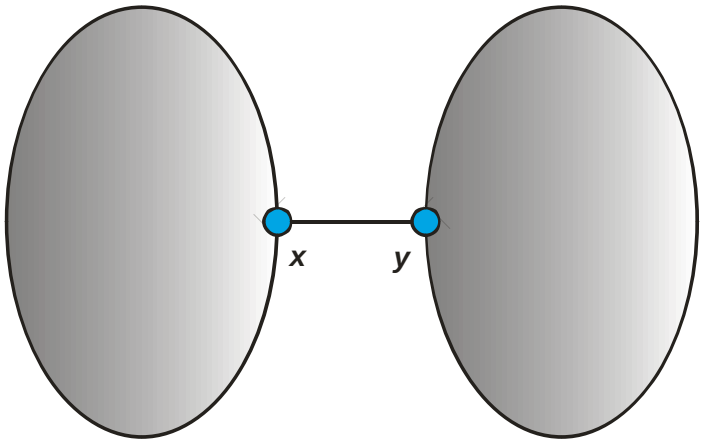


Figure 3-8. Nodes with expected high vertex betweenness B_x and B_y

In the context of technological networks, such as electric power or water distribution, the vertex betweenness can be measured for paths going from a generation subset $F \subseteq G$, to a distribution subset $H \subseteq G$ passing through a transmission subset $Q \subseteq G$ (Albert et al., 2004). Assuming that network flows are routed through the most direct path (i.e., shortest), which is consistent with optimal engineering systems design, the betweenness, B_v, of a vertex v represents the load or total flow that traverses it. Its calculation reduces to finding every shortest path from F to H and keeping the score of the nodes linked by each path. Rank ordering by vertex betweenness takes into account

information beyond network topology. It accounts for the potential flow patterns *within* the network. This parameter requires distinctions between vertices' roles and functions. Its investigation provides information about the effect on network performance induced by removing the nodes that are in between most other nodes. These nodes are not necessarily the most connected ones. Calculation of the vertex betweenness takes time $O(n^4)$. Conventional shortest path algorithms run in time $O(nm)$, recovering the path of the shortest paths is a separate problem that can run in time $O(n^3)$, and the algorithm needs to be run at most n times, 1 after every removal. Hence the total running time is $O(\ n\ (nm+n^3)\ ) = O(n^4)$.

3.3.3 Vertex transshipment rank-ordering

The concept of vertex *transshipment*, denoted W_v, is introduced by the author as an alternative measure to the vertex *betweenness,* B_v. As in the case of B_v, vertex transshipment also takes into account network flow patterns, but in a greater detail. W_v monitors the actual total amount of flow that passes through each vertex so that the flow demand at the distribution nodes is optimally satisfied by the supply flow sent from the generation nodes. This optimality amounts to solving not a linear problem but a convex minimum cost network flow problem. This kind of optimization minimizes the cost of routing flow from a generation subset F to a distribution subset Q, and assumes that the cost is a convex function of the amount of flow traversing the edges. Flow costs vary in a convex manner in numerous problems settings including: (1) power losses in an electrical network due to resistance, (2) congestion costs in a city transportation network, (3) expansion costs of a communication network, and (4) flow losses in water, gas or oil

networks due to pressure decay induced by leaks, breaks and terrain topography (Ahuja et al., 1993). This implies that in addition to network topology and partition of the network into subsets of specialized vertices, their net demands, traversal costs, and capacities should also be known. The net demands of the system are the difference between supply and demand flow quantities at every node. Costs are related to the price of traversing a particular edge, which is defined here to be proportional to the length of the edges. Capacity bounds—lower and upper—of every vertex and edge of the network refer to their physical constraints to allow flow traversal (i.e., pipe diameters, transmission line voltage limits, thermal insulation, age, etc.) The general formulation for an optimal convex cost network flow model, with directed edges having ends (i,j) oriented from vertex i to vertex j—or simply e_{ij}—is the following (Rardin, 2000):

$$\begin{aligned} \min \quad & \sum_{(i,j)\in E(G)} c_{i,j} x_{i,j}^2 \\ \text{s.t.} \quad & \sum_{(i,k)\in E(G)} x_{i,k} - \sum_{(k,j)\in E(G)} x_{k,j} = b_k \text{ for all } k \in V(G) \\ & l_{ij} \le x_{i,j} \le u_{i,j} \qquad \text{for all } (i,j) \in E(G) \end{aligned} \tag{3-8}$$

where $c_{i,j}$ represents the cost of traversing the edges connecting vertices (i,j); $x_{i,j}$ represents the decision variables of the model which are the optimal amount of flow that traverses edges $e_{ij} \in E(G)$; b_k is the net flow at each vertex; and $l_{i,j}$, $u_{i,j}$ are the lower and upper bounds of the flow within edges $e_{ij} \in E(G)$. Figure 3-9 presents a network of seven vertices, two of which are for generation, two for transmission and three for distribution (i.e., consumption). Also, the parameters required to solve a convex minimum cost flow problem

are identified along each edge. The minimum capacity of the edges is $l_{ij} = 0$ for all $e_{ij} \in E(G)$.

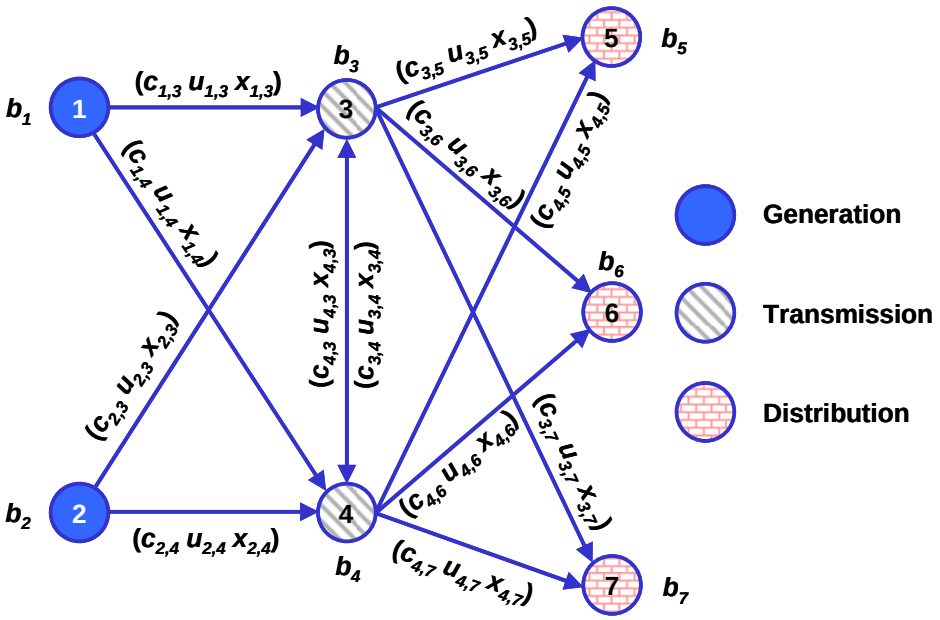


Figure 3-9. Parameters of generic convex minimum cost network flow problems

Once the optimal amount of flow $x_{i,j}$ to be routed through the edges is established, the vertex transshipment, W_v, is the summation of incoming flows at every vertex $v \in V(G)$. This quantity represents the importance of a node in terms of its role to maintaining optimal streams within the network. It provides a realistic assessment of the vertices that are essential for flow *circulation*. Also, since in real networks some vertices are operated close to their design capacities—allowing maximal flow circulation through them—there is a correlation between them and the vertices with high W_v. This advises that vertex transshipment rank ordering can also be used for determining effective mitigation actions whose benefits are expected to propagate optimally through the entire network.

This rank-ordering can be calculated with a polynomial-time algorithm. The worst-case complexity analysis indicates that W_v can be calculated in time $O(n^2m^2\log U)$, with U representing the edge with the largest flow upper bound. The minimum convex cost flow problem alone is solved in time $O(m \log U\, S(n, m, C))$, where $S(n, m, C)$ is the time required to solve a shortest path problem on a network with n vertices, m edges, and with C as the largest edge cost. The shortest path problem is solved in time $O(nm)$. Repeating the operation n times, once after every removal, leads to the indicated total time for generating the vertex transshipment rank-ordered set.

The three rank-ordering methods presented above allow devising vertex removal strategies that follow particular criteria for deciding vertex importance (i.e., vertex degree, vertex betweenness, and vertex transshipment). A final method for providing vertex removal sets can be implemented, and it has the characteristic that does not require any vertex input data.

3.3.4 No rank-ordering

In addition to rank ordering vertices according to their topological characteristics or according to their role in optimal flow circulation, it is important to consider failure of nodes due to aging, lack of maintenance, or unforeseen malfunction. Random vertex selection captures these events and complements the collection of rank-ordered vertex sets for implementation of element removal strategies. This strategy is also very efficient computationally since its complexity is linearly related to graph order n and hence its worst-case instance runs in time $O(n)$.

3.4 NETWORK PERFORMANCE MEASURES

Independent of the nature of the disruption and the mechanism by which it propagates—either by overloading the most connected vertices, or the vertices most in between other vertices, or the most instrumental to flow circulation, or by random malfunction—overall network functionality needs to be characterized. This characterization requires comparison of the response of the system after any disruption (i.e., network performance), with predefined network damage states. Response parameters could depend on network topology alone, or they can combine topology with various levels of refinement to capture flow patterns. Damage states can reflect different levels of functionality if they are set to be thresholds within the possible value range of the performance parameters.

Three network performance measures are introduced in this section. They, as in the case of the vertex rank-ordering strategies, are presented in increasing order of input information demand and relevance. Network response characterization, given its underlying architecture and layout for flow distribution, is essential to understanding failure modes, cascading mechanisms, growth patterns, and mitigation spread.

3.4.1 Efficiency, *E*

Network efficiency generalizes the concept of the *reciprocal harmonic mean*, *L'*, of a graph *G*. The fundamental difference is that *E* drops the restriction of edge *unweightedness* and *undirectedness*. As originally formulated, *L'* relies upon calculation of the distance $d(i,j)$ between any two vertices of the graph. The distance $d(i,j)$ was defined as the *length* of the shortest path *i-j*, and this length corresponds to the number of edges in the

path. Also, *L'* assumes the graphs to be undirected. The generalization introduces a matrix *D*, with non-zero entries identical to the adjacency matrix *A*, and d_{ij} elements that represent the physical distance between them. Hence, the new distances $d(i,j)$ calculated using *A* and *D* reflect the physical separation between vertices *i* and *j*, and $d(i,j) \geq d_{ij}$ for all $i, j \in V(G)$. The equality is valid if there is an edge joining *i* and *j*. The parameter *E* is calculated as follows (Latora and Marchiori, 2001; Latora and Marchiori, 2002):

$$E = \frac{\sum_{i \neq j} \frac{1}{d(i,j)}}{\sum_{i \neq j} \frac{1}{d_{ij}}} \qquad \textbf{(3-9)}$$

where the denominator normalizes the numerator with respect to the most efficient of all possible simple graphs: a *complete* graph with same order *n* as the original graph. Therefore, $0 \leq E(G) \leq 1$. This parameter, as its names suggests, can be regarded as a *global* indicator of efficiency in network connectivity. It represents the ease with which any two vertices can communicate or share flow. The input information requirements are limited to the Euclidean distances between adjacent vertices. Its calculation can be performed in polynomial time $O(nm)$. If calculated *n* times, once after every vertex removal, its running time grows to $O(n^2m)$.

3.4.2 Connectivity loss, C_L

Given a graph $G = \{ F \cup H \cup Q \}$, the concept of *connectivity loss*, C_L, is useful to quantify the average decrease of the ability of distribution vertices—those belonging to the

Q subset—to receive flow from the generation vertices—elements of the F subset—passing through transmission vertices—elements of the H subset. In other words C_L quantifies the decrease in the number of generators with connecting paths to the distribution vertices (Albert et al., 2004). Denoting n_F the order of the generation subset $F \subseteq G$ at the unperturbed state of the network (i.e., the number of connecting paths from every generation vertex to any distribution vertex), and n^i_F the number of generation units able to supply flow to distribution vertex i, after a perturbation takes place. The connectivity loss, C_L, can be calculated as:

$$C_L = 1 - \left\langle \frac{n_F^i}{n_F} \right\rangle_i \qquad \textbf{(3-10)}$$

where the averaging $\langle\rangle$ is done over the distribution vertices i of the distribution subset $Q \subseteq G$. The calculation of this parameter clearly relies on the topological structure of the network, and on the existence of paths connecting supply and demand elements. The input information requires distinction on the types of vertices, and as for the case of network efficiency, E, the paths are assumed to follow the shortest available route computed with Euclidean distances. The algorithm to compute C_L has a worst-case running time $O(nm)$. If the calculation is performed by systematically removing all vertices—according to any of the rank-ordering strategies of the previous section—the running time becomes $O(n^2m)$.

3.4.3 Service flow reduction, S_{FR}

Network efficiency, E, and connectivity loss, C_L, only provide information pertaining to the processes that take place within the network. However, it is of paramount importance to assess the effect that any network element removal has not only on internal network functionality, but also on the way it impacts its *end-users*. For that reason, the *service flow reduction*, S_{FR}, performance measure for generic networks is introduced in this study (Dueñas-Osorio et al., 2005a).

The objective in measuring S_{FR} is to quantify the amount of flow that does not meet the distribution vertex demands. These demands are directly related to the affected population, and therefore, it is more meaningful for decision-making and consequence minimization. The network service flow reduction, S_{FR}, can be calculated as:

$$S_{FR} = 1 - \left\langle \frac{S_i}{D_i} \right\rangle_{\mathrm{i}} \qquad \textbf{(3-11)}$$

where S_i denotes the actual amount of flow supplied to distribution vertex i at any given time, and D_i denotes the flow demand of vertex i which is considered to be proportional to the population served by distribution vertex i in the undisturbed network state. The averaging $\langle\rangle$ is done over the distribution vertices i of the distribution subset $Q \subseteq G$.

Calculation of S_{FR} requires a two-phase problem solving approach. The first phase solves a convex minimum cost network flow problem. In other words, phase I calculates the optimal amount of flow that traverses the network so that it meets the constraints imposed by matching supply and demand with minimum circulation costs. This phase is

performed even for *disconnected* graphs which may exist after network disturbances. Its calculation requires the use of infinite costs in the edges incident to any removed vertex. This condition ensures connectivity while providing optimal flows throughout the remaining network. The second phase of the problem solves a maximum flow – minimum cost problem. Essentially, phase II pushes as much available generation flow as the distribution vertices can consume. However, the upper bounds on the flow traversing each edge are set to be the optimal edge flows found in phase I.

This performance measure requires information about the amount of population served by each distribution vertex (i.e., demand), the supply capacity of the generation vertices, the cost of transporting flow (i.e., proportional to physical distances), and the maximum and minimum carrying capacity of the network elements. Despite of the apparent computationally intensive two-phase method to calculate S_{FR}, its insightful measurements can be performed in polynomial time. The worst-case complexity analysis indicates that it requires a running time $O(nm^2 \log U + n^2 m) = O(nm^2 \log U)$. This is dominated by the time of computing the convex minimum cost network flow. Calculating S_{FR} after every vertex removal, according to a rank vertex ordered set, the running time becomes $O(n^2 m^2 \log U)$. This parameter in addition to providing the link between network performance and potential social impacts also retains its computational feasibility for algorithmic implementation. Its uniqueness relies upon the combination of topological information with feasible optimal flow patterns.

There exist models for capturing network dynamics whose sophistication goes beyond topology and optimal flow. Such models utilize mass and energy conservation principles in the case of water networks, and they use circuit equations, in the case of

electric power (Hatestad, 2004; Advanced Grounding Concepts, 2005). However, their implementation is limited to specific systems, and the requirements for data imply knowledge of detailed sensitive information about *every* element of the networked system. For example, assessing the equilibrium of a power system requires for every vertex—or bus, as referred to in the electrical engineering community—specification of voltage magnitudes, reactive power, element resistance, and losses; and for every edge or transmission and distribution line, their capacities, lengths, and resistance. Information drawn from these models allows calculation of *exact* values that can be related with the impact of irregular network operation on end-users. However, S_{FR} is instrumental in providing, from a statistical point of view, the same information (i.e., perturbation impact on end users) at a mere fraction of the computation time and data requirements.

3.5 SUMMARY

Fundamental concepts developed in the field of graph theory have allowed important advances in network characterization. Several new properties have been proposed in this study to quantify the differences in network structure and flow patterns. Parameters such as mean distance, clustering coefficient, vertex degree distribution, and redundancy ratio are intended to provide statistical information about the configuration of the networks. The relevance that each network element has to maintaining connectivity and facilitating flow circulation is measured following various rank-ordering strategies. Vertex degree, vertex betweenness, vertex transshipment, and random selection are introduced as potential strategies to systematically remove network elements and investigate their effects on network functionality. This functionality can be measured

with different levels of detail depending on the network input information. Performance measures build up from the simplest parameter referred to as network efficiency, to a more elaborate parameter referred to as connectivity loss, to the most comprehensive, introduced here, referred to as the service flow reduction. The later takes into account topological properties of the network plus the patterns of flow that follow optimal paths within it, and yet it still retains computational feasibility for large networks. Optimality (i.e., minimum cost paths) negotiates realistic constraints such as flow transmission cost, flow capacity, and network element function. This parameter is expected to be more instrumental than other current statistical parameters in conveying the risk potential to decision makers and community leaders.

CHAPTER 4

NETWORK MODELS

Network modeling is essential to explaining properties observed in real complex systems. When a network model captures and reproduces measurable characteristics of graphs, it provides unique insights about the fundamental mechanisms that allow real networks to display particular properties. Interestingly, some properties have the character of *universality*. In other words, most networks possess properties dictated by similar underlying principles that guide their evolution and behavior.

The most primitive network model devised to capture complex real phenomena is the random network which is still used today as a reference system since it can be described by elegant mathematical theorems. This network is referred to as the ER model, in recognition to the developers Erdös and Rényi (Erdös and Rényi, 1959; Bollobás, 1998b).

More sophisticated models did not appear until the late 90's, when a very simple network referred to as the WS model—in recognition to their developers Watts and Strogatz (Watts and Strogatz, 1998; Watts, 1999)—was able to capture paradoxical properties observed in real social networks. Their model explained the mechanism by which tightly connected groups of acquaintances were also very close to apparently distant and unrelated groups. This phenomenon is referred to as the *small-world effect*.

Another breakthrough was made when certain real network properties, such as the vertex degree distribution, were better explained by models that displayed the small-world effect but contained realistic attributes like growth and preferential attachment. This dynamic model is referred to as the BA model in recognition to the developers Barabási and Albert (Albert et al., 2000; Barabási, 2003; Barabási and Bonabeau, 2003).

This chapter presents another type of network model with the purpose of realistically representing the class of technological networks that contain transmission and distribution, (TD), civil infrastructures. This model is intended to statistically characterize geographically distributed networks such as electric energy, water distribution, gas, and oil. Telecommunications networks have a different structure. They rely upon decentralized hubs for connectivity. Also, the geographical separation between elements is not as strong a constraint in as in the case of other transmission and distribution systems for global flow traversal. Calibration of the TD model relies upon analysis of data from real networks documented in published work. Two of these documented networks are examined in more detail because they share the same geographical region, and this makes them suitable for later interdependent analysis.

This chapter also presents a comparison of network properties including all artificial models and explores their variation as a function of network size. This is investigated to establish their *scaling* characteristics. The properties of real networks are also included. The chapter concludes with a discussion of potential robust architectures for effective resilience of distributed networks.

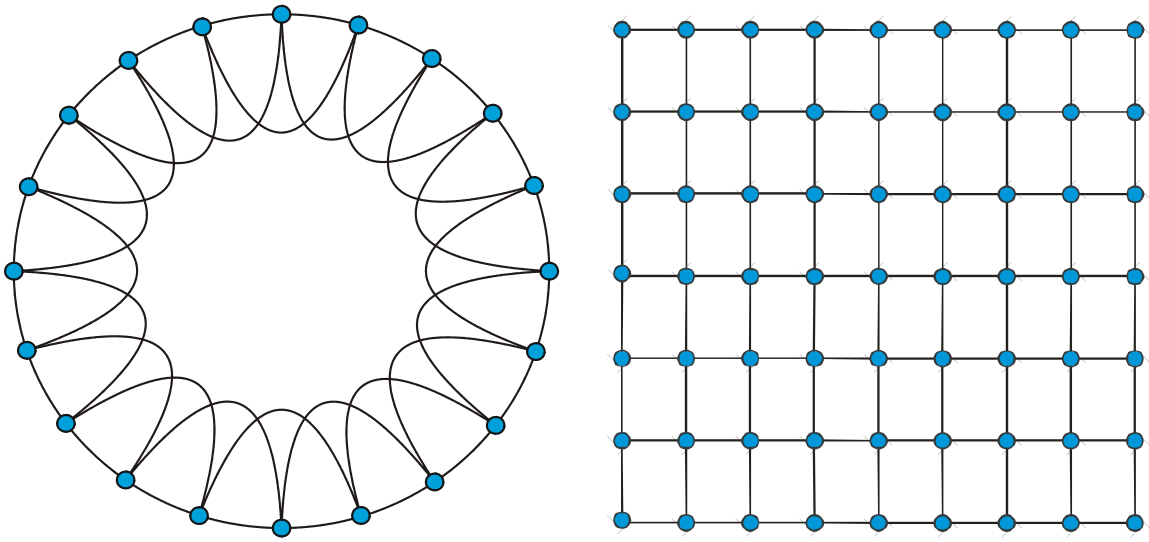
4.1 TRANSMISSION AND DISTRIBUTION NETWORKS

Generic civil infrastructure networks are designed to allow optimal circulation of flows (e.g., minimum cost, least resistance, etc.) from supply locations to demand locations. An intrinsic property of these networks is that their growth is correlated in location and number with the growth of the population that consumes the flow. Population growth at a particular location with fixed geographical extent has a saturation limit imposed by the density of the population, the availability of public services, and access to affordable space for living and working. This simply results in expansion of the original jurisdictional area from the existing core to its peripheries (i.e., radial growth). Network expansion follows this growth pattern. The result is a network with a mesh-like structure. This network topology exhibits differences according to the type of infrastructure. For a given geographical area, such as a county, oil and gas networks at the transmission level (i.e., main lines) are *sparse* networks. The power grid at the same transmission level is less sparse, and potable water is still less sparse than the others. Telecommunication networks are also sparse but do not display mesh-like structure. Rather, they evolve as decentralized topologies as shown in Figure 3-1.

An important property shared by all mesh-like topologies and pertaining to the ideal class of *lattice-graphs* is that they do not possess edges with *long range*—that is, edges connecting otherwise distant vertices. A lattice-graph, or more generally, a *d-lattice* is an unweighted, undirected, regular, simple graph of dimension d in which any vertex v is joined to its lattice neighbors, u_i and w_i, as specified by the following rules (Watts, 1999):

$$u_i = [(v - i^{d'}) + n] \pmod{n}$$
$$w_i = (v + i^{d'}) \pmod{n} \qquad \textbf{(4-1)}$$

where $1 \leq i \leq d(v)/2$, $1 \leq d' \leq d$, $d(v)$ is the vertex degree of v, which is equal for all $v \in V(G)$, n is the order of the graph G, and (mod n) is the arithmetic integer modulus that ensures *periodic* boundaries to the lattice. To illustrate this regular structures a 1-lattice with $d(v) = 2$ is a ring, and a 2-lattice with $d(v) = 4$ is a two dimensional square grid (Figure 4-1).



(a) (b)

Figure 4-1. Two d-lattices: (a) 1-latice with $d(v) = 4$, and (b) 2-lattice with $d(v) = 4$

The *range* of an edge e_{ij} denoted by $r(i,j)$ is the length of shortest path between i and j in the *absence* of that edge. An edge (i,j) is called an *r-edge* if it has a range $r(i,j) = r$; hence, the observation that geographically distributed networks do not possess long-range edges. This means that most infrastructure networks for flow transmission and

distribution cover vast extensions but with short range edges. This has a very important implication: the response and redundancy of a TD (transmission and distribution) network is largely dependent on the characteristics of the neighborhood of each vertex. Global network properties arise because of the characteristics exhibited by its vertices' neighborhoods, and neighborhoods of neighborhoods.

4.1.1 Construction of TD networks

In order to develop a flow TD (transmission and distribution) network, it is proposed to use modified 2-lattices with $d(v) = 8$ for all $v \in V(G)$ as *substrate* graphs. A *substrate* is a template graph topology. The modification consists of removing the *periodicity* of the boundary elements of the lattice. Therefore, vertices on the periphery of the graph will have smaller vertex degree than the internal vertices. This implies $d(v)_{\text{boundary}} < d(v)_{\text{internal}}$, which also implies that the average vertex degree, $d(G) < d(v)_{\text{internal}}$. However, for large n, $d(G) \sim d(v)_{\text{internal}}$. Figure 4-2 shows a single internal vertex of the 2-lattice with $d(v)_{\text{internal}} = 8$ substrate. This fundamental building block represents the simplest *bypass* mechanism. If a vertex fails, network flows can still circulate throughout the network without compromising global functionality by locally bypassing the malfunctioning element.

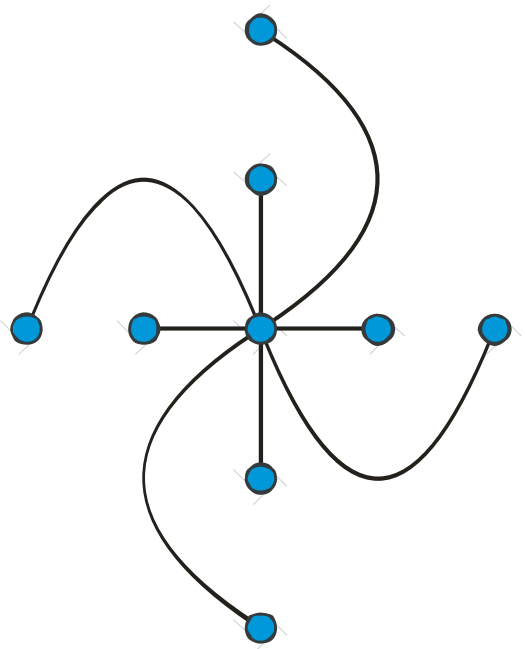


Figure 4-2. Building block of a TD model substrate: a vertex in a 2-lattice with $d(v) = 8$

A periodic TD substrate with $d(v) = 8$ for all $v \in V(G)$ and $n = 9$ is a complete graph. The number of edges in *any* periodic TD substrate is $m = ½\, n \times d(v)$. The number of edges for a specific aperiodic TD substrate with $d(v) = 4$ is $m = ½ [\, d(v)\, (n - n^{½}) \,]$, and with $d(v) = 8$ is $m = ½ [\, d(v)\, (n - {}^{3}/_{2} n^{½}) \,]$. For the cases in which the substrate is required to have *directed* edges, then m for periodic and aperiodic substrates is twice the number of the undirected cases. A directed substrate building block is depicted in Figure 4-3. Development of a TD model requires the following steps:

(i) Generate an undirected—or directed if required—aperiodic TD substrate of order n equal to the order of the real system, and with $d(G) \sim 8$.

(ii) Retain each of the m edges of the substrate with a probability of existence equal to p_m.

(iii) Repeat steps (i) and (ii) until an ensemble of $x \geq 20$ connected TD graphs is generated.

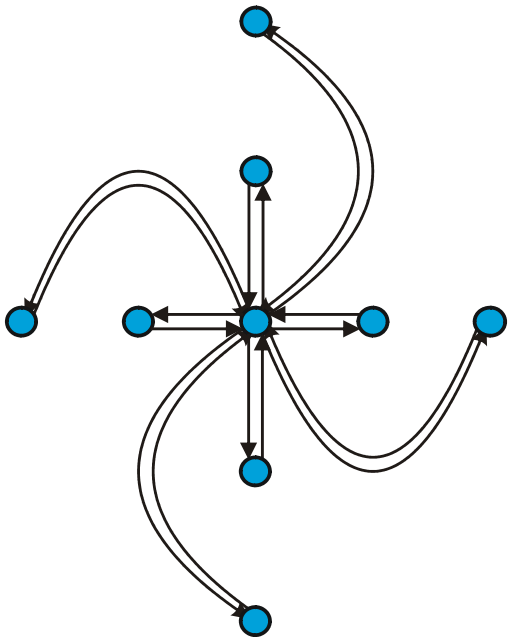


Figure 4-3. Building block of a directed TD substrate

The probability p_m is empirically set to be the ratio between the edge *density* of real infrastructure systems and the edge density of aperiodic TD substrates with $d(G) \sim 8$. Edge density is the ratio between the number of edges (i.e., nonzero entries of the adjacency matrix divided by two) of a graph G with n vertices, and the number of edges in a complete graph of same order n. When the networks need to be directed, these densities use the full set of nonzero entries in their adjacency matrices. Connectivity is assured by sequential retention of edges. Start at a vertex v, and retain one of its incident edges. Then, the end of the chosen edge provides the new vertex for repeating the step which is performed $p_m \times m_{TD\ substrate}$ times. At early stages, this allows populating the

adjacency matrix to form square grid networks, towards the end this procedure populates the matrix to form the bypasses typical of TD substrates.

Given the geographical constraints that prevent real transmission and distribution networks to have long range edges, it is expected that the edge density of TD substrates with $d(G) \sim 8$ exceeds the edge density of *most* measured real network edge densities. In the rare event that this condition is not met, the average vertex degree $d(G)$ of the TD substrate has to be increased. The upper limit for the average vertex degree is $d(G) \sim n - 1$. For periodic boundaries the equality holds and a simple complete graph is formed—the most redundant of all simple graphs.

Measurements of the number of edges m as a function of the order n for several real power networks are shown in Figure 4-4. Nine networks with different vertex set sizes are investigated from published work: (1) a network with $n = 14{,}099$ that represents the major elements of the North American power grid (Albert et al., 2004); (2) a network with $n = 4{,}961$ that represents the power grid of the states west of the Rocky Mountains (Watts, 1999); (3) a network with $n = 59$ representing the major components of the power grid of Shelby County, TN (Multidisciplinary Center for Earthquake Engineering Research, 1998); and (4) six typical networks with $n = 9$, 14, 30, 39, 57, and 118, used as simulation tool by researchers and educators (Zimmerman et al., 2005). For comparison purposes, Figure 4-4 also presents the scaling of m as a function of n for aperiodic TD substrates with $d(G) \sim 8$, and for the simple complete graphs with $d(v) = n - 1$.

Power law fitting of the measured *scaling* of m as a function of n for power grids provides the following empirical relation:

$$m = 1.2n^{1.02} \quad \textbf{(4-2)}$$

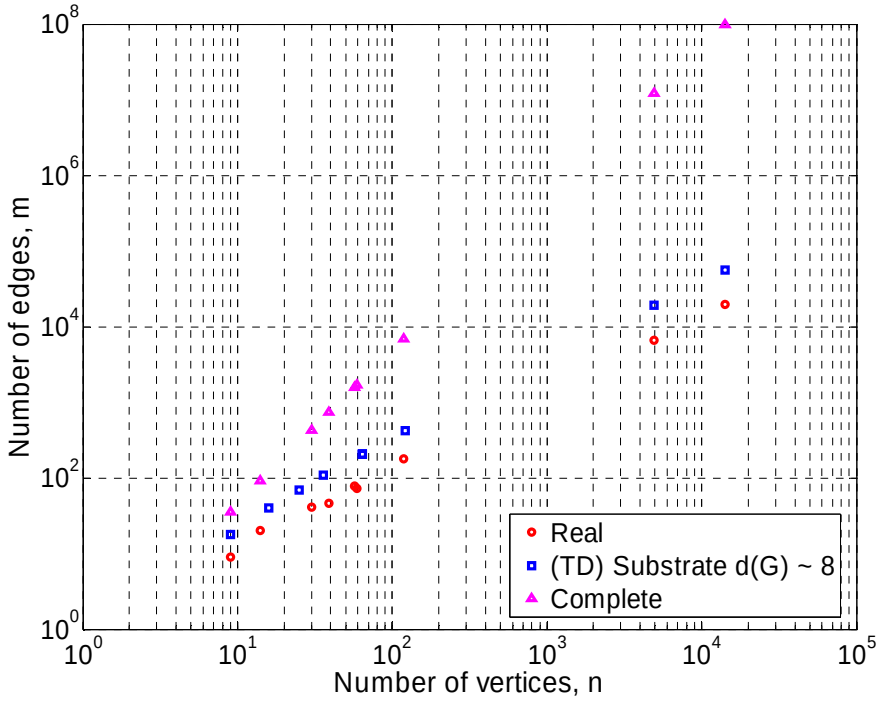


Figure 4-4. Power law dependence between m and n

The exponent for the empirical relation of complete graphs on n is twice as large as the measured for real graphs. The TD substrates with $d(G) \sim 8$ have a slightly larger exponent as compared to the real power grids. The empirical equations for these mathematically constructed graphs are:

$$m = 2.18n^{1.07}, \quad \text{for TD substrates with } d(G) \approx 8 \quad \textbf{(4-3)}$$

$$m = 0.46n^{2.01}, \quad \text{for simple complete graphs} \quad \textbf{(4-4)}$$

These empirical relations provide the means to estimate the probability of edge retention, p_m, in power grids with respect to TD substrates. The ratio between densities of real graphs and TD substrates reduces to simply calculating the ratio between their number of edges, m. The scaling of p_m for power grids—as a particular case of transmission and distribution networks—is:

$$p_m = 0.55n^{-0.05} \quad \textbf{(4-5)}$$

where $n \geq 1$, and p_m indicates the proportion of edges that are retained from a TD substrate. The particular substrate used so far (i.e., modified 2-lattice with $d(G) \sim 8$), recognizes the limitations in edge range, and provides the means to capture absorption of disturbances at local scales. Figure 4-5 shows the scaling of the number of edges, m, as a function of n, for the graphs presented in Figure 4-4, and adds the scaling for graphs that correspond to power grid realizations of the TD substrate.

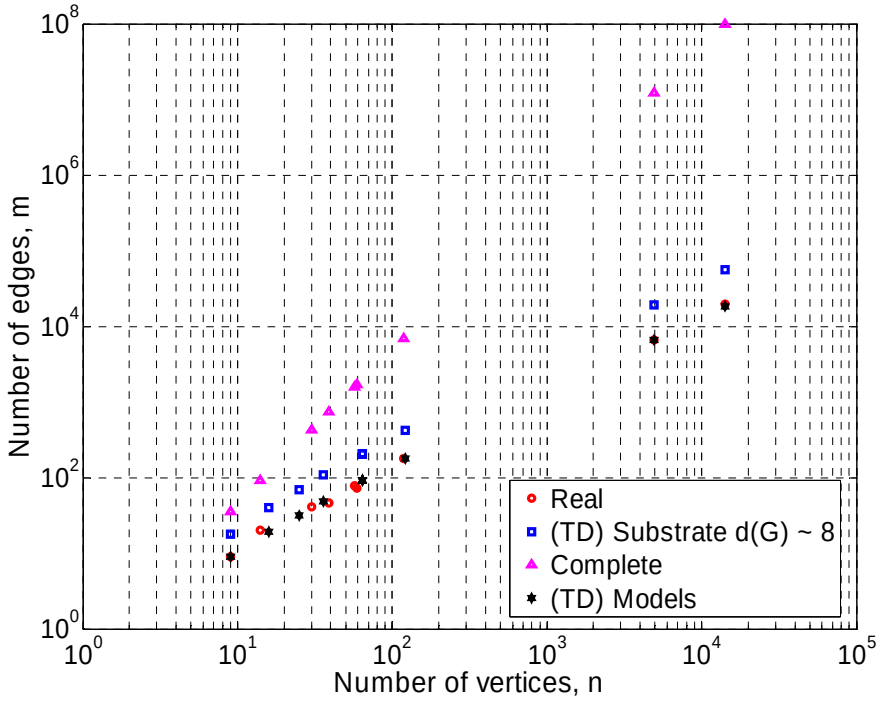


Figure 4-5. Scaling of edge count in TD models for power grids

Generation of TD models calibrated for other geographically distributed networks such as potable water, natural gas, or oil, proceeds in a similar way to the development of TD models for electric energy. Their scaling of m with respect to n is assumed to have the same functional power-law form $m = \alpha n^{\beta}$. This is because the number of edges that a new vertex adds to m is relatively constant due to geographical constraints, and therefore it grows at a proportional rate with n, given similar exponents β. In other words, as n becomes larger, the number of edges per vertex, $d(v)$ is confined to be as large as the order of the union of its i^{th} neighborhoods for small i—in this discussion $i = 1, 2$. This implies that $d(G)$ experiences small variations as n increases. Hence, the differences between TD models are reduced to the proportionality constant α. TD substrates with $d(G) \sim 8$ have $\alpha < 4$, square grids (i.e., aperiodic 2-lattices with $d(v) = 4$) have $\alpha < 2$,

trees and aperiodic rings have $\alpha < 1$, and totally fragmented networks—those with the number of components equal to n—have $\alpha = 0$.

The probabilities of edge existence p_m for ideal mathematical models, empirical power grids, and other arbitrary TD models (e.g., water, gas, and oil) are presented in Table 4-1. Water and gas network probabilities of edge existence are assumed to be approximately ±10% the value of power grids. Oil networks are less connected and are assumed to be 20% less dense than power grids. The probability values for ideal lattice graphs is obtained with respect to periodic TD substrates (i.e., $d(G) = 8$). For square grids, open rings, and utility networks, the probabilities are obtained with respect to aperiodic TD substrates (i.e., $d(G) \sim 8$).

Table 4-1. Edge existence probabilities, p_m, for networks $G \subseteq$ TD substrates

Network Type	Probability of edge existence, p_m	
	With respect to *periodic* substrates	With respect to *aperiodic* substrates
2-lattice with $d(v) = 4$	0.50	_
1-lattice with $d(v) = 2$	0.25	_
Square grid	_	$0.65n^{-0.03}$
Open ring	_	$0.42n^{-0.06}$
Empirical power	_	$0.55n^{-0.05}$
Speculated water	_	$0.60n^{-0.05}$
Speculated gas	_	$0.50n^{-0.05}$
Speculated oil	_	$0.45n^{-0.05}$

It is clear that the aperiodic TD model for power grids lies between square grids and open rings. The speculated values for the other lifelines try to reflect their inherent

topological characteristics. Oil networks tend to be minimally connected, while water distribution networks tend to be mesh-like connected. Gas networks, tend to be more cyclic than the oil infrastructure, but its widespread use is heterogeneous, making them sparse meshes. Figure 4-6 illustrates the limiting topologies of TD models.

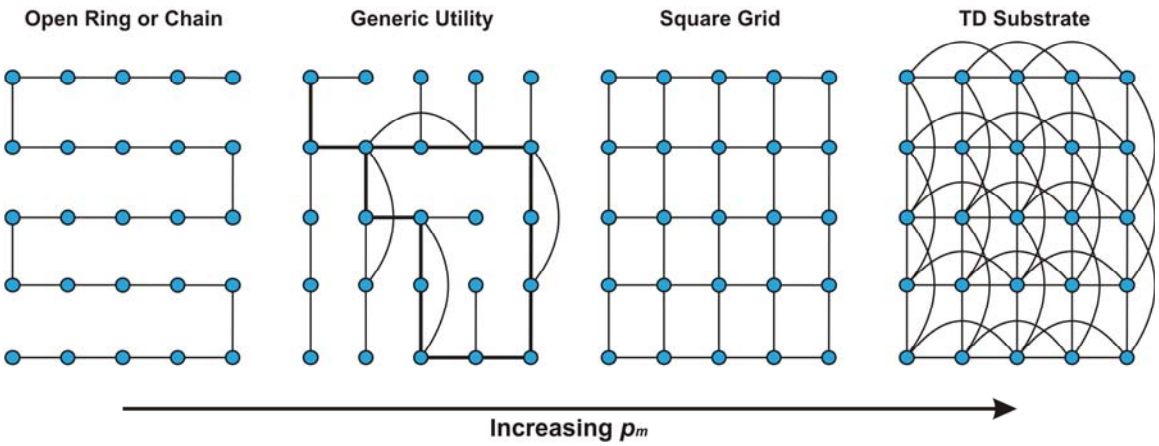


Figure 4-6. Sample of topology realizations for the TD models as a function of p_m

4.2 REAL POWER AND WATER NETWORKS

Network model development is useful for the analysis of network properties and the prediction of network response to disturbances. However, understanding network characteristics has an ulterior purpose: to aid in explaining the effects of interactions among them if their coupling is acknowledged. For this reason, two networks will first be introduced independently and described in particular detail. They are chosen because their spatial distribution shares the same geographical area. The networks are located in the Central United States in Shelby County, Tennessee, on the east banks of the Mississippi River. Their intended function is representative of critical infrastructures for

modern society. The networks are an electric power network, denoted by P, and a water distribution network, denoted by W (Chang et al., 1996; French and Jia, 1997; Hwang et al., 1998; Multidisciplinary Center for Earthquake Engineering Research, 1998). The water network is simplified so that the distribution system covers the entire area, but portions of its dense end-user delivery grid are removed from consideration. It just retains enough links to reasonably represent arterial water mains and major secondary feeders. The order of the systems are $n_p = 59$ and $n_w = 49$ for the power and water networks, respectively. The order of the generation node subsets $F_p \subseteq P$ and $F_w \subseteq W$ is $n_{Fp} = 8$, and $N_{Fw} = 15$.

The water network topology resembles a two dimensional grid with quadrilaterals and few acyclic branches as fundamental building blocks or network *motifs*. The power grid exhibits sparser quadrilaterals and more acyclic branches. This is proven numerically by their average *clustering coefficient*, γ, and average *vertex degree*, $d(G)$. These parameters estimate the average proportion in which the neighbors of vertex v are adjacent to each other, and the average number of links per vertex, respectively. In this case $\gamma_p = 0.034$, and $\gamma_w = 0.047$, although, both values indicate sparseness, there are *fewer* local clusters in the power network. Regarding their vertex degree, $d(P) = 2.48$, and $d(W) = 2.90$. This indicates the tendency of water nodes to have more than two connections. A value of $d(G) = 2.67$ is reported for the 4,961 vertices of the United States Western power Interconnection (Watts, 1999). Figure 4-7 shows the physical layout of the power grid, and Figure 4-8 shows the layout of the water distribution network.

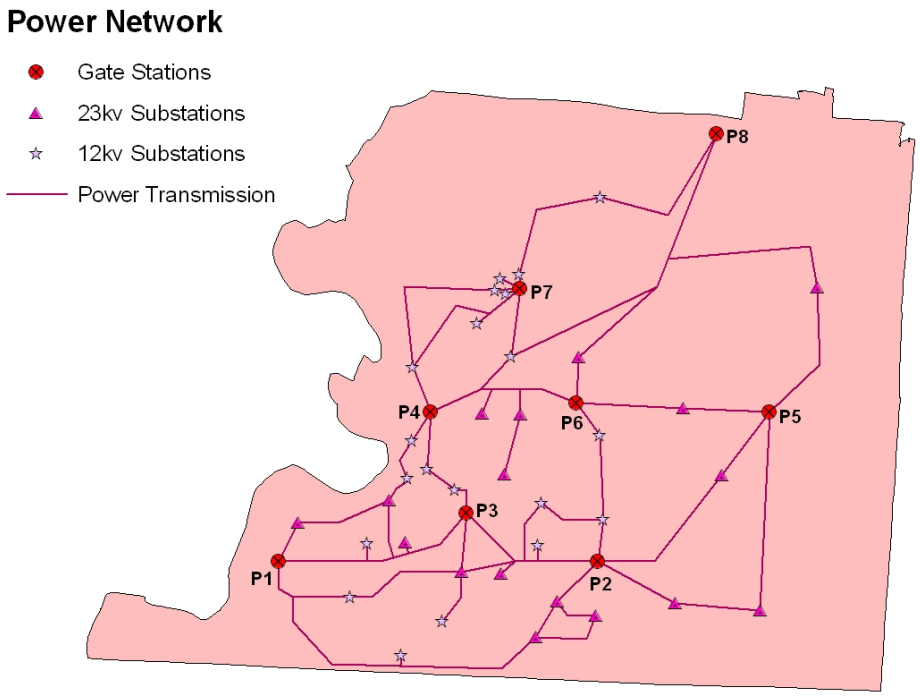


Figure 4-7. Simplified power grid on 59 vertices for Shelby County, Tennessee

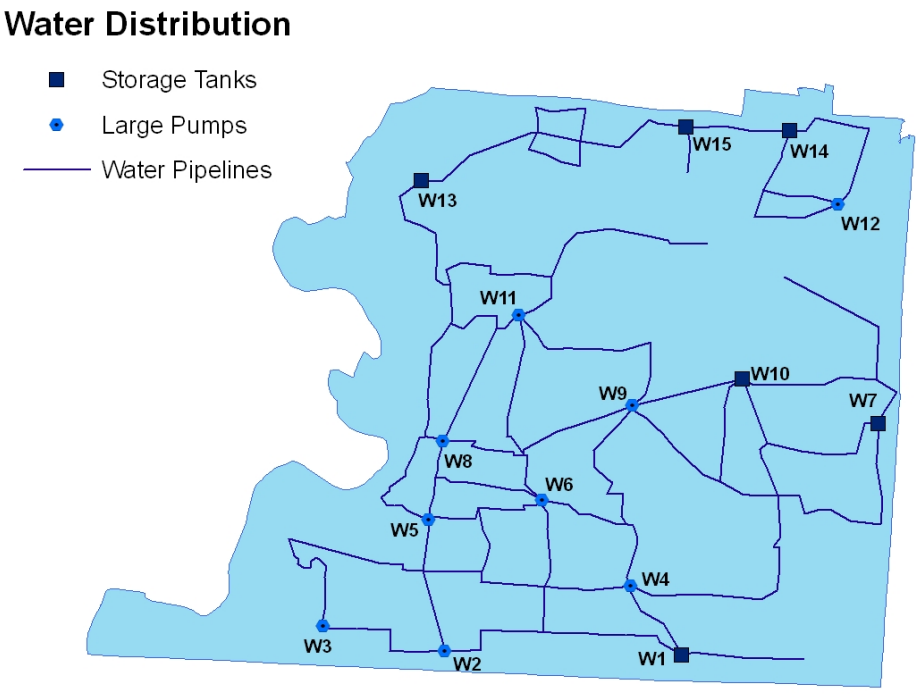


Figure 4-8. Simplified water network on 49 vertices for Shelby County, Tennessee

A summary of the fundamental topological properties of these two networks is presented in Table 4-2. Five mathematical graphs are included for comparison: two graph realizations of the TD model (i.e., one for power and one for water), one TD substrate, one square grid, and a complete graph. The trend is clear: the more the number of edges the more desirable the properties exhibited by the graphs. The number edges, m, for the mathematical graphs increases from TD model realizations, to square grids, to TD substrates, to complete graphs. An increasing number of edges produce graphs with more cycles, which essentially provide alternative routes for flow traversal. This increment in cycle existence directly impacts the measured properties in the networks. Large m increases the reciprocal harmonic mean, L', because it is easier to connect any vertex v to any vertex w. More edges increase the average vertex degree, $d(G)$, and also increase the redundancy ratio, R_R, due to the increment in alternative independent paths between nodes.

The two real networks exhibit similar properties. They both exist in a Euclidean plane, but the water network displays slightly better metrics due to its tendency to resemble square grids. The TD models representing power and water networks maintain the same relative differences of the real systems. These TD models (i.e., realizations of the TD substrate) exist on squares adjacency matrices. Therefore, their properties are reported for two orders: $n = 49$ and $n = 64$, which contain the order of the real systems. Their measured properties include the values displayed by the real infrastructures. The variability in their properties is small except for the clustering coefficient, γ. This variability is expected because the TD model realization occurs on a template that

provides the means to either form or not node *triads*—direct connections among neighbors of a vertex *v*. Square grids do not form triads, hence $\gamma = 0$.

Table 4-2. Topological properties of networks with *n* ~ order of real systems

Network Type	Fundamental Network Properties			
	Reciprocal harmonic mean, *L'*	**Vertex degree *d*(*G*)**	**Clustering coefficient, γ**	**Redundancy ratio, R_R**
Real 59-node Memphis Power	0.251	2.474	0.034	0.188
Real 49-node Memphis Water	0.264	2.898	0.047	0.193
TD model with power p_m on 49 (64) nodes	0.271 ±0.009 (0.238 ±0.007)	2.857 ±0 (2.906 ±0)	0.023 ±0.018 (0.021 ±0.019)	0.199 ±0.009 (0.197 ±0.008)
TD model with water p_m on 49 (64) nodes	0.289 ± 0.009 (0.258 ±0.003)	3.102 ± 0 (3.156 ±0)	0.044 ± 0.018 (0.038 ±0.028)	0.212 ± 0.010 (0.213 ±0.010)
TD substrate with *d*(*G*) ~ 8 on 49 (64) nodes	0.440 (0.399)	6.288 (6.500)	0.261 (0.255)	0.295 (0.288)
Square grid on 49 (64) nodes	0.301 (0.267)	3.428 (3.500)	0 (0)	0.221 (0.221)
Complete graph on 49 (64) vertices	1 (1)	48 (63)	1 (1)	1 (1)

An interesting observation from the previous table is that the measured properties for the same systems, but with different *n*, exhibited noticeable variations. Exploration of these network property variations as a function of their order, *n*, is referred more

generally to as the *network parameter scaling* which was initially introduced to measure the variation of m as a function of n.

4.3 NETWORK PARAMETER SCALING

The scaling of network properties as a function of graph order, n, provides valuable insights about the effects of topology on measured network characteristics. These scaling trends are the basis for empirical relationships or analytical expressions—based on counting arguments—to characterize network properties as their size and order grows larger. For large n it becomes impractical to implement computer algorithms to measure network properties. In typical contemporary personal computers—processor speed ~ 3 GHz and RAM ~ 1GB—handling of operations involving large adjacency matrices for network characterization imposes virtual memory limitations. Therefore, scaling analysis enables exploration of topological properties in large graphs.

The topologies considered in this scaling exploration include random ER graphs, WS graphs, BA graphs, square grids, TD substrates, TD models of power and water systems, real power grids, and complete graphs. The order of the networks ranges from $10 \leq n \leq 200$.

The scaling of m, L', $d(G)$, $\Delta(G)$, γ, and R_R as a function of n is shown in the figures below. All graphs are constructed so that they have similar order, n, and size, m, except for the cases in which their inherent topology requires a larger size (e.g., complete graphs or TD substrates). For the networks that are realizations of particular classes (e.g., TD models, WS models, ER models, and BA models), their metrics are an average of 20 simulations.

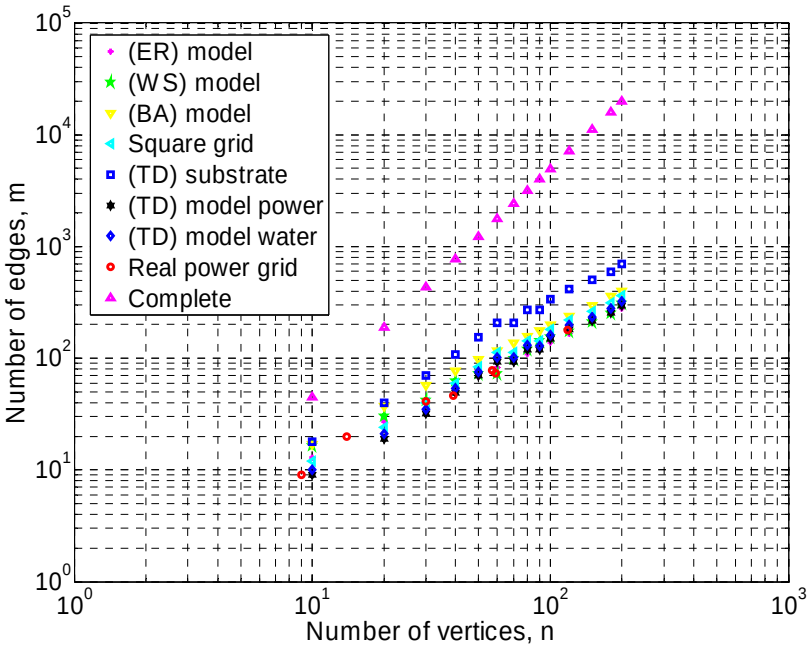


Figure 4-9. Scaling of the number of edges, *m*

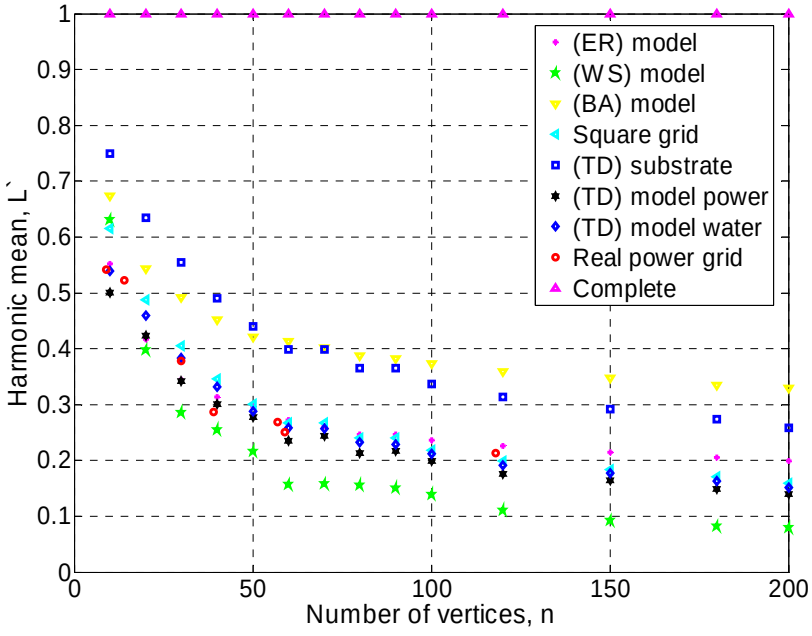


Figure 4-10. Scaling of the reciprocal harmonic mean, *L'*

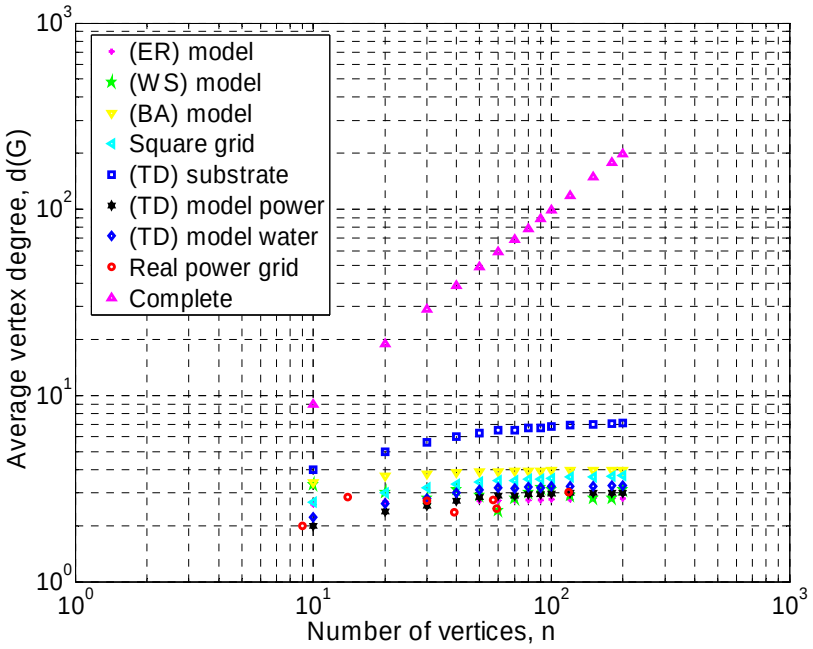


Figure 4-11. Scaling of the average vertex degree, $d(G)$

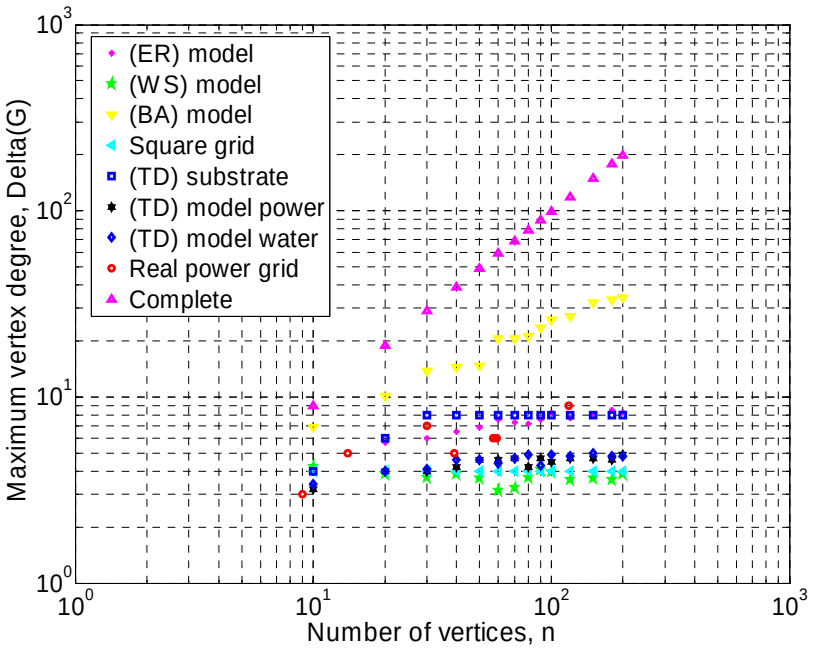


Figure 4-12. Scaling of the maximum vertex degree, $\Delta(G)$

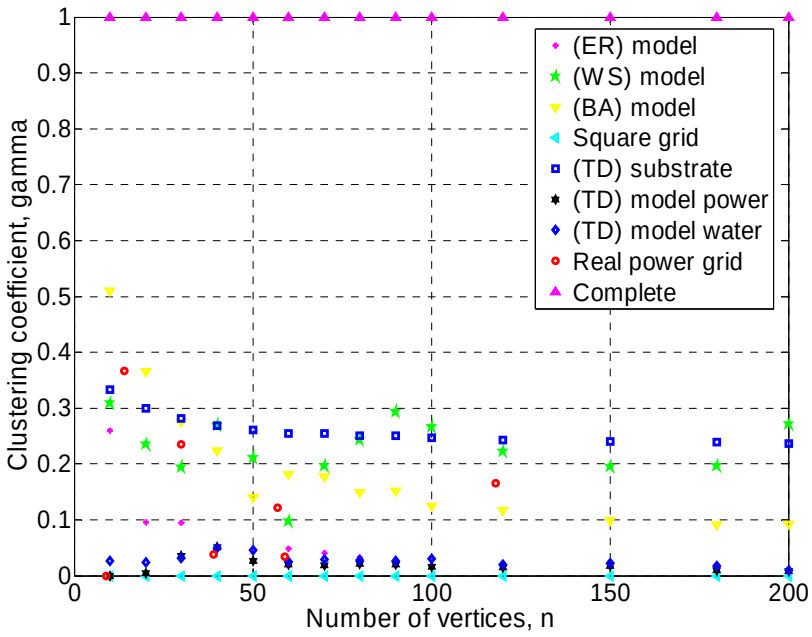


Figure 4-13. Scaling of the clustering coefficient, γ

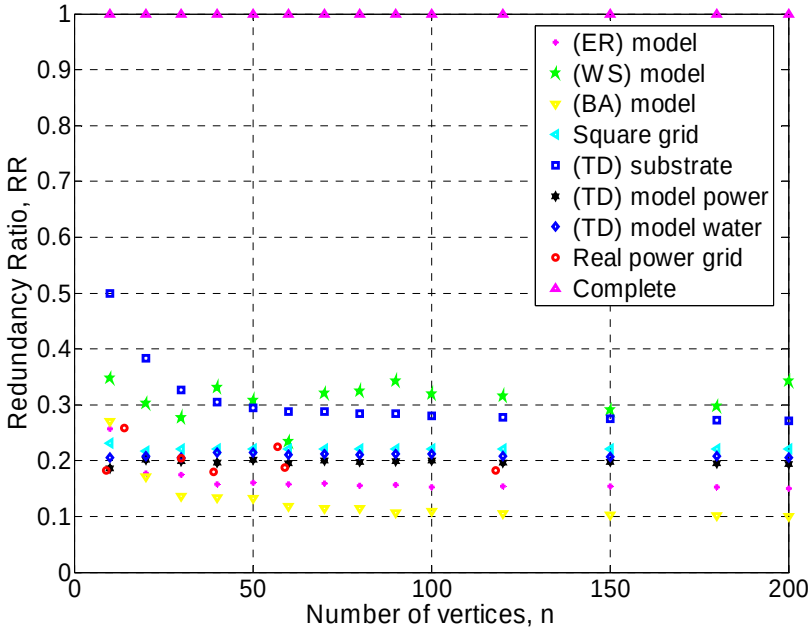


Figure 4-14. Scaling of the redundancy ratio, R_R

Figure 4-9 revisits the scaling of m. Except for the mathematical constructs of complete graphs and TD substrates, the rest of the models have a comparable number of edges. This is important because assessment of other network properties will only be function of n and the topology of the graphs. Table C-1 in the appendices contains a summary of the analytical and empirical relations of the scaling of m for mathematical (i.e., artificial), and real networks. The expressions are obtained from counting arguments and power law fitting.

Figure 4-10 shows the scaling of the reciprocal harmonic mean, L'. This parameter measures the potential for efficient global connectivity. This potential is a slow decreasing function for large n. The realizations of the TD substrate compare well with each other, with real graphs, and with square grids. The slowest connections are provided by the WS model because of its ring-like structure, which sets every node apart from others as n increases. The BA model shows significant L' due to its few hub nodes.

Figure 4-11 and Figure 4-12 deal with vertex degrees. The average vertex degree, $d(G)$ for the TD substrate and the square grids approaches the degree of perfect 2–lattices with $d(G) = 8$ and $d(G) = 4$. The TD models exhibit slightly smaller degree than the square grids, in agreement with measured values for large real power grids with $d(G) \sim 3$. Complete graphs have a linear scaling relation of $d(G)$ with n. The maximum vertex degree, $\Delta(G)$, highlights the existence of greatly connected nodes in BA models. The other TD models have larger maximum degree as compared with square grids. This is because the realizations of TD models from TD substrates allows for the existence of bypasses, as displayed by real networks.

Figure 4-13 displays the scaling of network clustering coefficients, γ. This parameter, as a measure of the probability of neighbor connectivity at the local scales, exhibits small variation for the TD models. Square grids prevent neighbors of a vertex, v, to be directly connected, and therefore the clustering coefficient remains null. Mathematical constructs such as complete graphs, WS models and TD substrates show no significant variation for $n > 50$. The clustering coefficient, γ, in ER and BA models tend to zero with large n because there is an increasingly small probability of randomly connecting the neighbors of vertices, or of connecting new vertices with low degree nodes.

Figure 4-14 presents the scaling of network redundancy ratio, R_R. This quantifies the ability of individual nodes to reroute flow in case their immediate neighbors are nonfunctional. For all graphs, this parameter remains constant as a function of n, for $n > 25$. Distributed networks (e.g., TD models and square grids) show similar redundancy and perform better than simple random graphs. Their values are bounded below by the BA model, whose local cooperation is minimal due to their growth with preferential attachment to highly connected vertices. The general constancy of the R_R is due to the modular structure that each graph has as it grows. They use similar building blocks for expansion, thus maintaining similar local capabilities for rerouting flow. The variation of the parameter is also small. Figure 4-15 shows that network realizations of a particular class (e.g., TD models, BA models, and ER models) exhibit low standard deviation. For large n, R_R shows a coefficient of variation (COV) ~ 0.05. The WS models display larger variation because they are constructed as a realization of 1-lattices with random rewiring of their vertices. 1-lattices require even vertex degree. In these simulations the average

degree for distributed networks is $d(G) \sim 3$. This odd value forces realizations of the WS models to have $d(G) = 2$ and $d(G) = 4$, thus increasing the variability. Real graphs and mathematical constructs such as complete graphs, TD substrates, and square grids have no variation in their property characterization because of their deterministic nature. The results for TD models, WS models, ER models, and BA models are an average of 20 simulations.

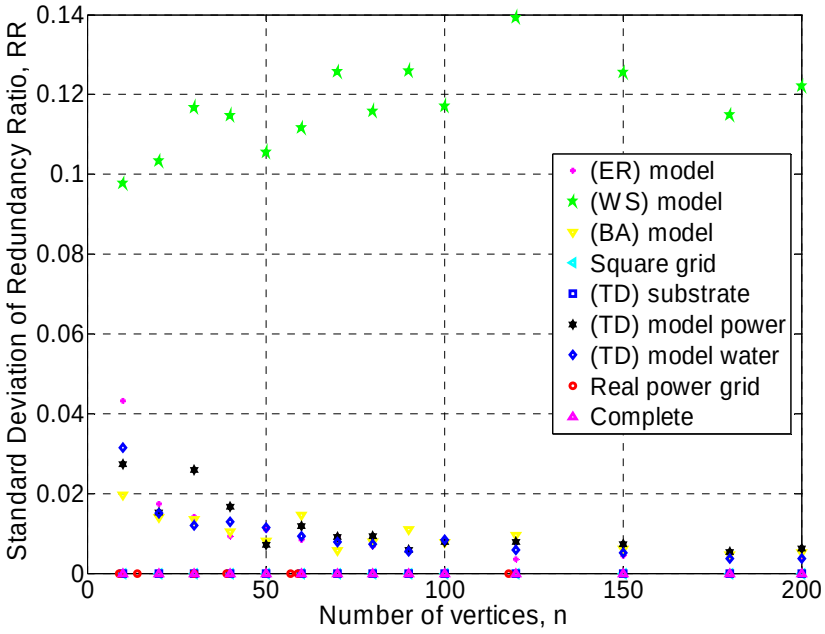


Figure 4-15. Standard deviation scaling of the redundancy ratio, R_R.

4.4 SUMMARY

Network models have successfully allowed researchers to explain the mechanisms by which real networks evolve and display specific topological properties. Random, small-world, and scale-free networks are just a few examples of graphs whose rules for

development explain properties of several real networks thus giving them a character of *universality*. In the context of infrastructure networks, the most common generic type of network is that for minimum cost flow transmission and distribution of goods or services between predefined locations. This motivated the development of a new template graph model that contains most of the geographically distributed lifelines. The template model is referred to as a TD substrate which provides bypasses to all vertices. Boundary elements have slightly less possibilities for bypassing, but for large n the properties of internal vertices dominate the measured properties of these networks. These TD substrates are used to generate TD models of other lifelines. A tuning probability of edge existence allows realizations of these utility networks, whose properties compare well with real power networks.

Two specific networks are described in detail because of their usefulness for later interdependent analysis. The properties of these real networks (e.g., power and water), which share the same geopolitical extent, are compared with mathematical constructs of similar characteristics and with the realizations of the TD substrates. Comparisons showed agreement between real systems and their modeled counterparts. Also, some sensitivity of the properties was observed as a function of n. This prompted the exploration of the relationship between measured properties and the order, n, of the graphs. Such scaling analysis revealed that most networks experience a power law decrease in their global connectivity effectiveness. It also showed an almost constant local connectivity, γ, and constant redundancy ratio, R_R, which measures the ability of the flow to bypass malfunctioning nodes or edges. The modular growth patterns exhibited by

artificial and real networks ensure this constancy. Such modularity also determines a small COV for the redundancy of the networks.

Other topologies not explored in this work but that are part of ongoing investigations include decentralized models for telecommunication networks, and a very robust ideal network for distributed systems: plane triangulations. These graphs exhibit desirable properties and are not required to be complete. They just need to be maximally plane, or their vertices must be elements of graph triangles (i.e., cycles of length three or C^3). More precisely, a plane graph G is called a plane triangulation if every face of G—including the outer face—is bounded by a triangle. Figure 4-16 shows an idealized plane triangulation, and a modified triangulation for geographically distributed systems. This modified version violates the definition of plane graphs, where edges do not intersect with each other. However, it is useful to depict the possible topology of triangulation within distributed networked systems and to highlight its redundancy potential. Edge intersections of the drawing do not necessarily exist in real systems.

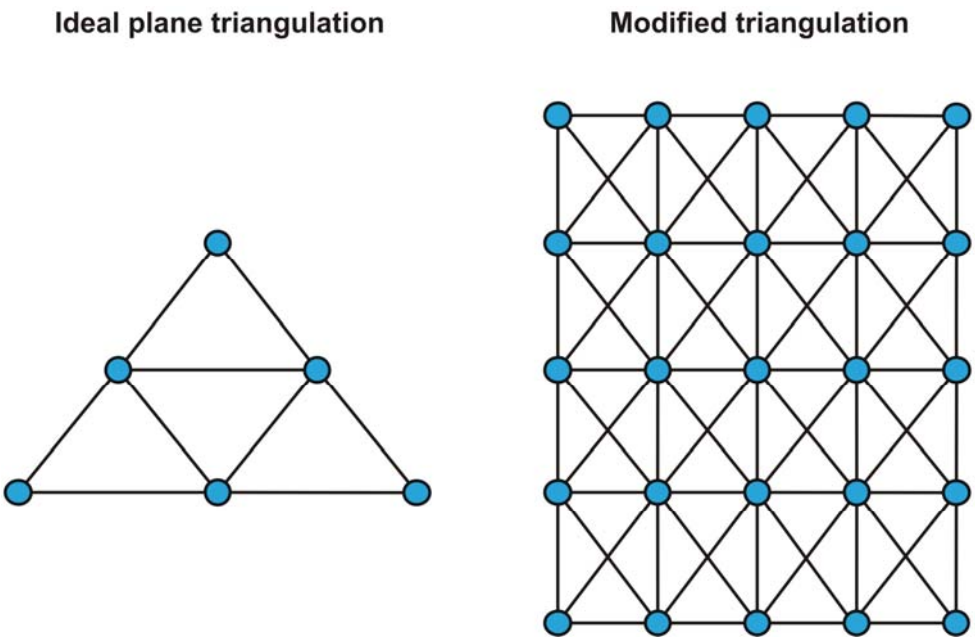


Figure 4-16. Plane triangulations as potential block for robust distributed networks

It is conjectured that future highly redundant systems will tend to form maximally plane graphs. One of the reasons is their obvious ability to locally absorb disruptions without the need to involve multiple neighborhoods per vertex as the existence of bypasses requires. Another reason is because end-users in this contemporary digital-quality era will demand exceptionally reliable flow supply which will necessarily depend on efficient redundancy mechanisms for networked systems. Finally, physical boundaries for the expansion of networks will prevent the traditional replication of building blocks. This implies that growth will head towards densification of the networks, thus increasing the number edges. Plane triangulations are efficient for increases in graph size, m, if confined by spatial boundaries.

CHAPTER 5

MODELING OF NETWORK INTERDEPENDENCIES

It is currently accepted that understanding the effect of interacting forces among infrastructure networks is of paramount importance for ensuring the continuous functionality of modern societies. Infrastructure interdependencies materialize in several ways, and systems are frequently connected at multiple points through a variety of mechanisms. There exist feedback and feedforward physical as well as nonphysical paths that create intricate branching topologies and make network interdependency increasingly intractable.

Even though there have been significant conceptual and theoretical advances in the field of interdependent networks, most frameworks still use highly simplified real networks, or use ideal network models to formulate the interdependencies problem. These frameworks do not explicitly question the role of network topology and network elements in maintaining and facilitating inter-infrastructure connectivity and flow exchange. Network topological and flow diffusion characteristics can be used in devising policies for minimization of the impact of disturbances on network functionality.

In order to proceed further with modeling interdependencies, this chapter reviews the conceptual *dimensions* that are currently accepted for describing infrastructure interdependencies. Then, it proposes a set of rules that allow objective characterization

of multiple-system interconnections. The application of these rules is illustrated with the two real systems (i.e., electric energy and potable water for Shelby County, TN) introduced in the previous chapter.

5.1 DIMENSIONS OF INTERDEPENDENCIES

For a given a geographically bounded system, determination of what constitutes interdependencies among infrastructures is a nontrivial problem because factors of different nature contribute to the observed coupling characteristics. Ideally, interdependencies could be quantified by weighting each of the following dimensions: type of interdependency, coupling and response behavior, infrastructure characteristics, infrastructure environment, type of failure, and network state of operation (Rinaldi et al., 2001). A brief description of these dimensions is presented below.

5.1.1 Types of interdependency

There are four main types of interdependencies: physical, geographical, informational, and logical. Physical interdependencies arise when the state of the infrastructures depends on tangible or material linkages among them. This type of interdependency is suitable for quantification with low uncertainty. Visual inspection would be the simplest procedure for determining physical coupling. Geographical interdependencies are also quantifiable because they depend on the spatial proximity among network elements. Informational interdependencies emerge due to the contemporary need for exchange of data, and the computerized control of infrastructure

performance. Finally, logical interdependencies appear when human decisions intervene, or when the connection is not physical, geographical, or informational.

5.1.2 Coupling and response behavior

This dimension is characterized by the degree of coupling, the coupling order, and the complexity of the connections. The degree of coupling refers to the "tightness or looseness" of the interdependency. In highly coupled infrastructures disturbances tend to propagate rapidly through and across them. Hence, tight coupling implies time dependent processes with little slack.

Coupling order defines whether or not two infrastructures are directly connected to one another or indirectly coupled through one or more intervening infrastructures. For instance, if the power grid fails, there is a supply / demand imbalance whose first-order effects could be disruption in gas supply, oil extraction, and water distribution. Second-order effects manifest as reduced cogeneration for oil production, inventory buildup in refineries, inventory drawdown in storage terminals, and crop losses in agriculture. Third-order effects can also be quantified by reduced oil production, shortages of gasoline for road transportation, disruption of flight schedules, financial losses, etc. This type of analysis can be extended to n^{th}-order effects, at the expense of introducing additional complexity and uncertainty.

Complexity of the links represents the linearity or nonlinearity of the interactions among critical infrastructures.

5.1.3 Infrastructure characteristics

This dimension considers the scale of the infrastructure in spatial and temporal aspects and the hierarchy of the elements that constitute an infrastructure. Scale brings into consideration the size of the problem analysis (i.e., level of granularity).

Different temporal scales for network functionality are illustrated in the case of unforeseen network element perturbation: in power grids disruption travels in milliseconds; in gas, water and transportation networks, disruptions propagate in hours; and in most networked systems, interruptions due to maintenance, upgrading, or capacity increase, propagate disruptive effects during years.

Spatial scales reflect the level in the hierarchy of network elements, at which the analysis is defined. Typical hierarchical levels, illustrated with a power grid, are:

- **part:** smallest component identifiable in power analysis.
- **unit:** functionally related collection of parts (e.g., steam generator).
- **subsystem:** array of units (e.g., secondary cooling system).
- **system:** grouping of subsystems (e.g., nuclear power plant).
- **infrastructure:** complete collection of like systems (e.g., electric power infrastructure).

A network *G*, as represented in this research by its adjacency matrix, is a collection of systems. The functionality of infrastructures is established from the performance of systems, subsystems, units, and parts.

5.1.4 Infrastructure environment

This dimension refers to the impact on network interconnectedness due to economic concerns, government decisions, legal regulations, public health and safety, technical limitations, and social and political decisions. These types of factors are highly subjective, and can introduce large uncertainties in the definition of network interdependencies.

5.1.5 Types of failures

Failure modes influence interdependent behavior through the way networks propagate or absorb disruptive effects. Cascading, escalating, and parallel failures are feasible modes of failure. Cascading failures are a manifestation of the potential vulnerability of otherwise highly robust networks. Blackouts are examples of such propagation mode, where disturbances can initiate and propagate in a variety of ways that are difficult to foresee. Escalating failures occur when an existing disruption in one infrastructure exacerbates an independent disruption of a second infrastructure, generally in the form of increasing the severity or the time for recovery or restoration.

A parallel failure occurs when two or more infrastructure networks are disrupted at the same time and elements within each network fail because of some common triggering event (e.g., earthquakes, cyclones, etc.). This case represents simultaneous effects, either because the elements occupy the same physical space, or because the root problem is widespread.

5.1.6 State of operation

The state of operation of an infrastructure exhibits different behaviors during normal operating conditions (i.e, varying from peak to off-peak conditions), during times of severe stress or disruption, and during times when repair and restoration activities are underway. The state of operation of a unit, subsystem, or system at the time of failure affects the extent and duration of the disruption of the entire infrastructure.

Even though the six dimensions for describing infrastructure interdependencies are conceptually simple, the amount of data required for their characterization makes their systematic evaluation impractical. Fewer fundamental parameters are necessary to capture the essential mechanism for coupling networked systems. In this study, three aspects are selected for establishing the model of network interdependencies. Accessibility to data, objectivity, and possibility to explore their domain are the criteria for their selection.

5.2 FUNDAMENTAL INTERDEPENDENCIES

Real networked systems require continuous multi-network interactions to perform their intended functions (National Research Council, 2002). Power grids depend on gas networks to fuel generation units. Water networks provide cooling and help controlling emissions of coal-based power generators. Water and gas networks are heavily dependent on power for pumping stations and control systems. Therefore, characterization of network performance needs to be enhanced by explicitly accounting for their interconnectedness. In

this study the approach for establishing network interactions is based upon geographical immediacy, direction of the interaction, and strength of the coupling.

A set of conditional probabilities of failure is established to describe the direction and strength of the coupling between adjacent multi-network elements. Considering two networks, for example the power and water networks of Shelby County, TN, their directed degree of coupling can be expressed as follows:

$$P(\text{Failure } W_j | \text{Failure } P_i) = p_{W_j|P_i}, \quad \text{for all } j \sim i \qquad \textbf{(5-1)}$$

where W_j represents failure of the j^{th} element of the water network; P_i is failure of the i^{th} element of the power network; $p_{Wj|Pi}$ is the value of the conditional probability of failure of element W_j given failure of element P_i; and ~ indicates node adjacency between every W_j element and every P_i element. The conditional probability, $P(W_j|P_i)$, can be tuned to represent any strength of network coupling, ranging from independence $P(W_j|P_i) = P(W_j)$ to complete interdependence $P(W_j|P_i) = 1$. Interdependent adjacencies are established from the geographical proximity between network elements.

Transmission and distribution networks provide service to their end-users through service areas assigned to specific distribution vertices. The size of the service areas is proportional to the population density and the capacity of the distribution nodes. An illustration of the use of geographical proximity for establishing interdependencies is shown in Figure 5-1. For clarity this figure only contains 23kV and 12kV substations, storage tanks, and large pumping plants. The highlighted area is served by one 23kV electrical substation that transforms power to its end users. Among its users there is a critical large

pumping station that fills the tank in the service area, and contributes to three storage tanks in the borders. In general, one substation provides power to one assigned service area. In this power network there are 37 service areas. Each service area is an aggregation of administrative and political subdivisions referred to as *census tracts*. The population size of the power service areas is on average 25,000 people. As an example, if power substation $P_{25} \subseteq P$ fails, it conditions with probability $P_{Wj|P25}$ the failure of pumping station W_{12}, and storage tanks W_7,W_{10}, W_{14}, W_{15}, for j = 7, 10, 12, 14, 15.

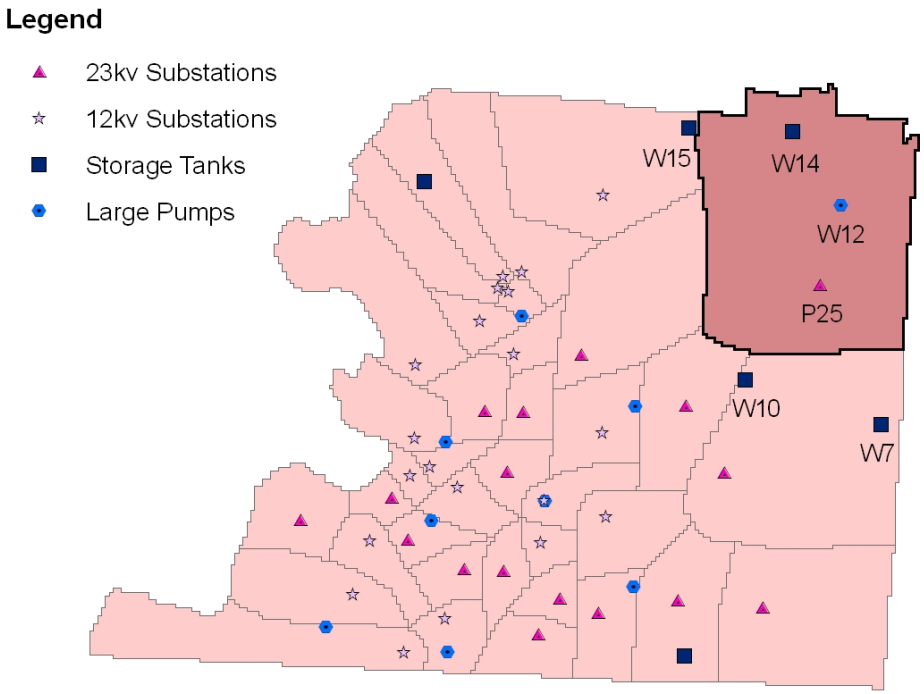


Figure 5-1. Distribution of electric power service areas

Failure of node P_{25} implies that approximately 9% of the county population, according to the 1990 United States census data (United States Census Bureau, 2002), has the potential to experience impaired water delivery service. Interdependent effects have the

potential to leave the water system incapable of fulfilling basic objectives: meeting water needs, accommodating fire fighting, and equalizing grid operating pressures.

Location, direction, and strength of the interdependencies are captured by the interdependent adjacency matrix, $A_{W|P}$. In this example $A_{W|P}$ is a $n_P \times n_W$ matrix whose non-zero entries, row by row, are the j^{th} water network elements served by the i^{th} power element. The values of these no-zero entries are the conditional probabilities $p_{Wj|Pi}$. For example, the 25th row of the matrix has five non-zero entries: columns 7th, 10th, 12th, 14th, and 15th. The values of the non-zero entries are the tunable strength of coupling $p_{Wj|P25}$ for j = 7, 10, 12, 14, 15. The $n_W \times n_P$ interdependent adjacency matrix, $A_{P|W}$, which captures the effect of water network element failure on power grid performance, is established in identical manner. Figure 5-2 displays 34 service areas for the potable water network in Shelby County, TN. These areas are defined for the *junctions* of main artery lines and secondary feeders of the simplified real water distribution network.

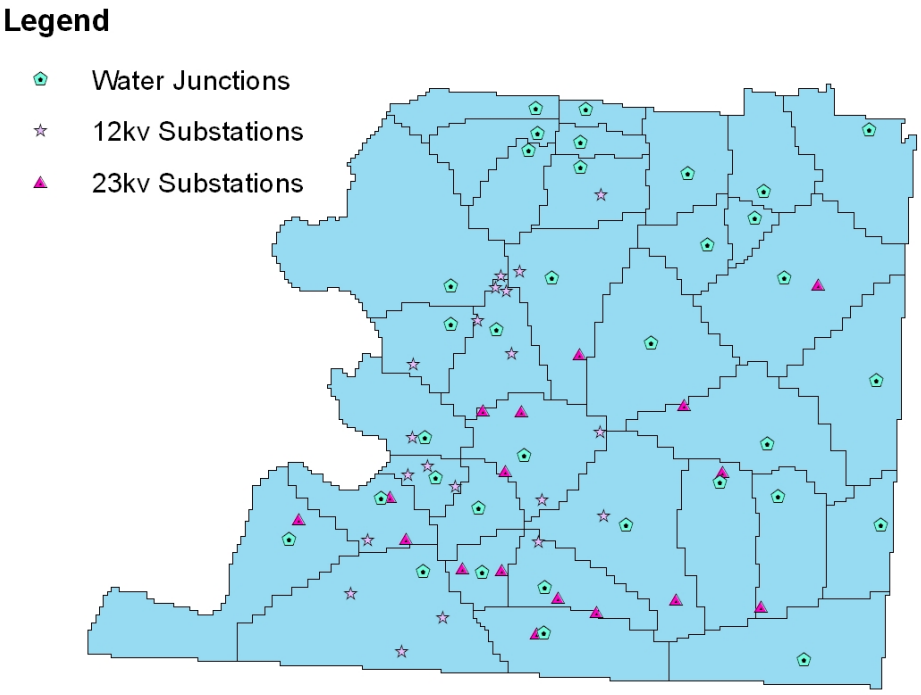


Figure 5-2. Distribution of potable water service areas

5.3 SUMMARY

Six conceptual dimensions for describing infrastructure interdependencies are considered representative of the mechanisms that trigger interconnectedness. Although these dimensions (i.e., type of interdependency, coupling and response behavior, infrastructure characteristics, infrastructure environment, type of failure, and network state of operation) are intuitive and conceptually simple, their quantification is limited by data availability. This study proposes exploration of more objective and fundamental aspects to characterize the interconnection among networked systems: location, direction, and strength of coupling. These fundamental aspects are contained in the interdependent adjacency matrices, $A_{G|G'}$, of interacting systems which is constructed from the conditional

probabilities $P(G_j|G'_i)$. The nonzero entries of these matrices indicate the location of the interconnection and the strength of the coupling, $p_{Wj|Pi}$, which is defined as a tunable variable to allow exploration of coupling from independence to complete interdependence. The location of the nonzero entries is obtained from geographical immediacy analysis of the interacting systems. The tunable parameter represents the conditional probability of failure of one infrastructure element given failure of distinct infrastructure element. This simple approach improves on the most recent schemes utilized for modeling interdependencies among networks with specific topological features.

CHAPTER 6

NETWORK RESPONSE TO DISTURBANCES

One of the persistent questions in the analysis of the performance of networked systems is what would happen to their functionality if particular network elements are removed? To answer this question it is necessary to measure the post-disturbance performance of the networks in terms of their capacity to remain functional or to fulfill customers' demands. *Generic* network performance measures that are meaningful to different infrastructure systems are desirable metrics. Critical infrastructures are often the type of networks that operate transmission and distribution of flows by either transfer or replication. Their performance can be measured using either simple metrics that only depend on the topology of the network, or metrics that require a combination of data from topology and flow patterns. These more elaborate metrics require knowledge of the role that each element has in the supply-demand chain, in addition to their flow capacities and cost of flow traversal as measured in terms of geographical separation. In addition, more detailed performance measures developed for specific networks may have extensive input data requirements necessary to model time-dependent physical processes. This specificity currently makes impractical their implementation for the analysis of *interdependent* network performance.

This chapter reviews two of the most recent proposals for generic network performance quantification—network efficiency, E, and network connectivity loss, C_L (Latora and Marchiori, 2001; Albert et al., 2004). These generic metrics are suitable for statistical exploration of network response to disturbances. Networks can be either real systems, simplifications of real systems, or mathematical constructs. In addition, this chapter uses the proposed network service flow reduction, S_{FR}, as a generic performance measure that captures network topology, flow patterns, social impact, and yet is still computationally efficient and suitable for statistical exploration of networked systems.

This chapter also illustrates the utilization of network performance measures when the systems are subjected to systematic removal of their elements. Removal of vertices is executed according to three strategies that reflect their importance to connectivity, their ability to facilitate flow, and their actual role in providing optimal flow distribution. This systematic removal is first performed on independent systems, and then its effects are investigated in interconnected networks.

This chapter includes a discussion of network response to seismic hazards as well. These events are some of the most disruptive and uncertain external disturbances and require a probabilistic framework for their investigation. The study of seismic hazard potential to disrupt network functionality is extended, based on a review of the seismic vulnerability of network elements, to the effect of their failure on infrastructure functionality. The parameters chosen to characterize network performance are the ones that combine topology and flow patterns (i.e., C_L, and S_{FR}). Finally, this analysis is expanded to the response of interdependent infrastructures. The chapter concludes with the introduction

of interdependent fragility curves as a means to characterize the probabilistic response of interacting infrastructures.

6.1 SYSTEMATIC DISRUPTION OF NETWORKS

Systematic network element removal refers to the process of deletion of vertices following a predefined policy. This policy dictates the order in which nodes have to be removed from the networks. Node rank-ordering is updated after every removal because deletion of vertices changes the topology and conditions the flow traversal within the networks.

This study considers four removal strategies to span several criteria that define the importance of nodes and the nature of the disruptive event (e.g., targeted attack, overload, random malfunctioning, etc.). Vertex degree rank-ordering is based on network connectivity and only requires topological information of the systems. Vertex betweenness rank-ordering requires information of the role that each node plays in the chain of flow generation, transmission, and distribution. Its calculation combines topology and introduces essential elements of optimal flow. Transshipment flow rank-ordering is the most comprehensive criterion and requires additional information to characterize the network as a minimum convex cost flow problem. This implies knowledge of the flow capacities of the elements—both nodes and edges—and the cost of pushing flow through particular paths based on Euclidean distances. Finally, random selection of vertices complements the set of removal policies. This strategy is included because it captures events such as lightning, vehicle collisions, falling trees, equipment failure, vandalism, animals, and planned

removals. These events are frequent in geographically distributed infrastructures, and they induce constant small disruptions which in some cases can develop into cascading failures.

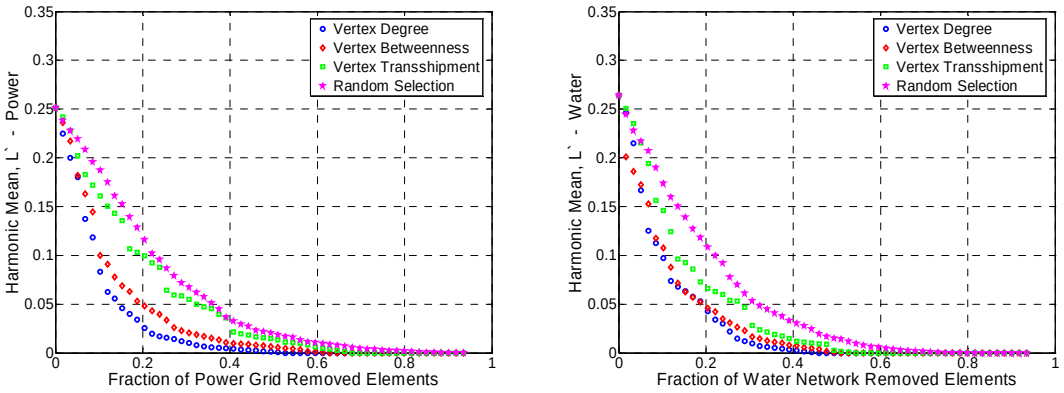
6.1.1 Independent networks

Exploration of independent network performance as a function of the fraction of removed rank-ordered vertices constitutes the fundamental step for response characterization. The objective is to identify the ways in which different networks degrade their functionality as the number of removed vertices increases, and also as the removal strategy varies.

This section presents results for simplified models of the power grid and water network of Shelby County, TN. The performance metrics for characterization of the disturbed networks are: global efficiency of the network, represented by its reciprocal harmonic mean, L'; average connectivity loss, C_L; and average service flow reduction, S_{FR}. The response of the networks for each performance measure is recorded for four different element removal strategies. These vertex removal strategies systematically delete nodes according to their importance in terms of: vertex degree, vertex betweenness, vertex transshipment flow, and random selection. The importance of each node is reassessed after every removal because of changes in topology and flow patterns. Chapter 3 contains definitions for each of these network performance measures and vertex removal strategies.

Figure 6-1 contains the response of the networks to different disruption mechanisms in terms of their reciprocal harmonic mean, L'. This parameter measures the ease for global connectivity in terms of their topology. Both networks display a similar trend in their response. This is expected since both systems have topologies that resemble regular square

grids. A revealing result is that the response to the various removal strategies is bounded by two mechanisms: the random selections of nodes, and the selection of vertices according to their vertex degree, $d(v)$. The first one captures unforeseen malfunctioning of network elements (e.g., lighting, vandalism, aging, collisions, animals, etc.), while the second one captures highly intrusive disruptions focused on the most topologically connected nodes. This vertex degree rank-ordering provides an idea of network response if malicious attacks were planned against them.



(a) Power grid (b) Water network

Figure 6-1. Reciprocal harmonic mean, *L'*, for a power grid and a water network

The removal strategies that are based upon vertex flow distribution within the network are contained between the random selection and targeted removal bounds. These flow-based removal strategies better reflect the possible evolution of a cascading failure. If an overloaded node fails (i.e., the one with the most flow traversing it), then its flow has to be redistributed. The next node more likely to fail is such node that after flow redistribution gets closer to, or exceeds, its maximum capacity. Figure 6-2 presents a normalized version

of the reciprocal harmonic mean, *L'*. This plot shows their response for the two observed bounds.

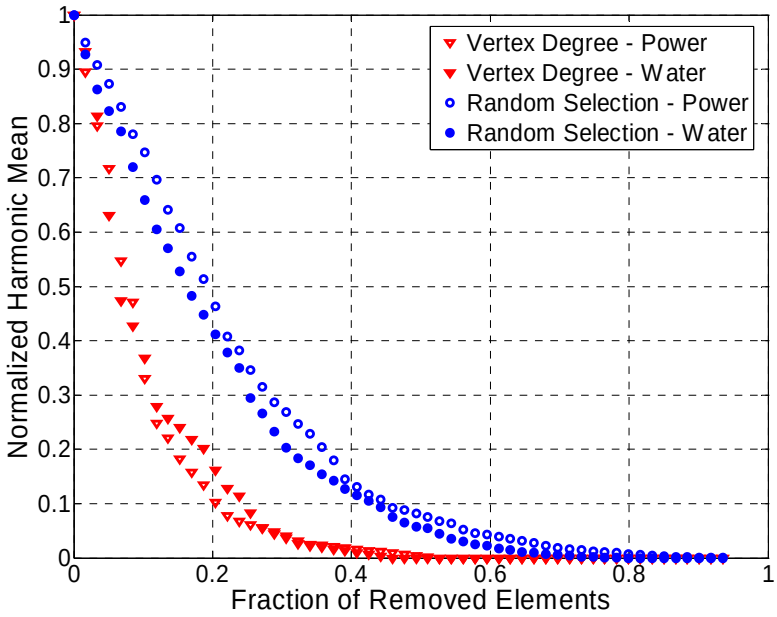
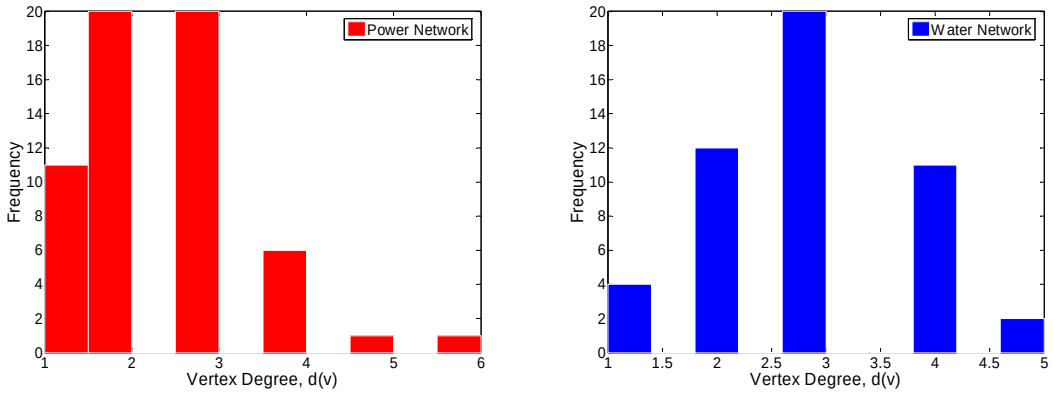


Figure 6-2. Normalized reciprocal harmonic mean *L'*

It is observed that for random selection, the water network performs poorer than the power grid. This is because the water network has properties closer to the square grids—its vertex degree exhibits a Poisson distribution—where every node is a *typical* node. Hence, removal of any water node counts since each node has a similar role in facilitating global connectivity. In the power grid this does not hold. There are some *special* nodes (e.g., more connected than the average node), which are difficult to pick at random. However, the price is that for targeted removals, the power grid becomes more vulnerable. This paradox is rooted in their topological characteristics. Figure 6-3 presents their vertex degree distribution. The power grid has a coefficient of variation,

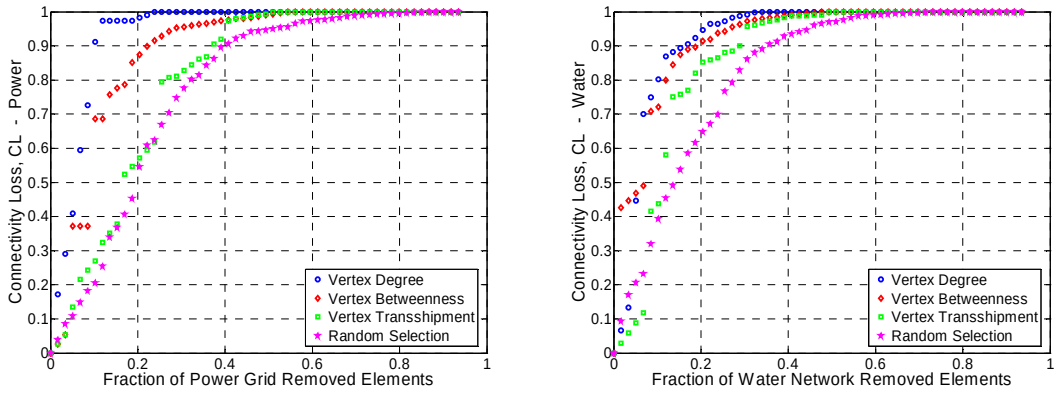
$COV_p = 0.44$, and a skewness coefficient, $\phi_p = 0.66$. The parameters for the water network are $COV_w = 0.34$ and skewness coefficient, $\phi_w = -0.05$—almost symmetric. The larger skewness of the power grid indicates the existence of a few highly connected that if strategically removed can induce significantly reduce network functionality.



(a) Power grid (b) Water network

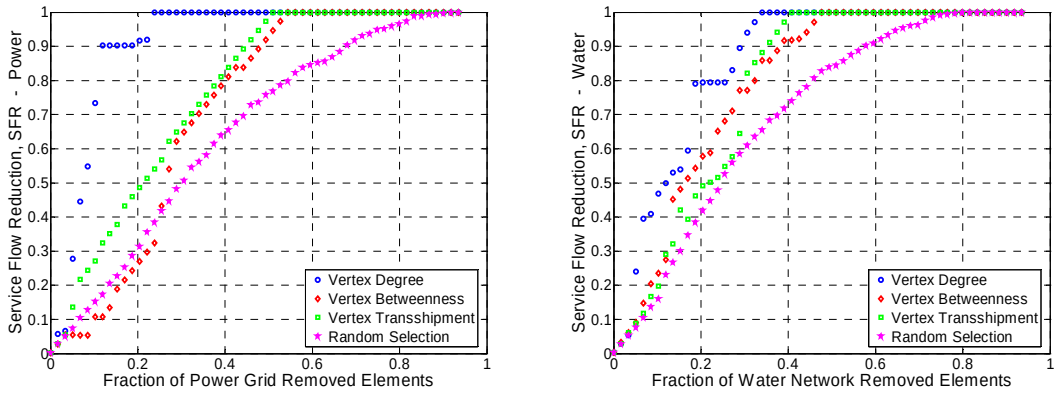
Figure 6-3. Vertex degree distribution for a power grid and water network

The other two network performance measures follow the same trends displayed by the global reciprocal harmonic mean, *L'*. However, C_L (Figure 6-4) and S_{FR} (Figure 6-5) provide a more realistic representation of network performance after disruption because they account for topology and flow traversal.



(a) Power grid (b) Water network

Figure 6-4. Connectivity loss, C_L, for a power grid and a water network

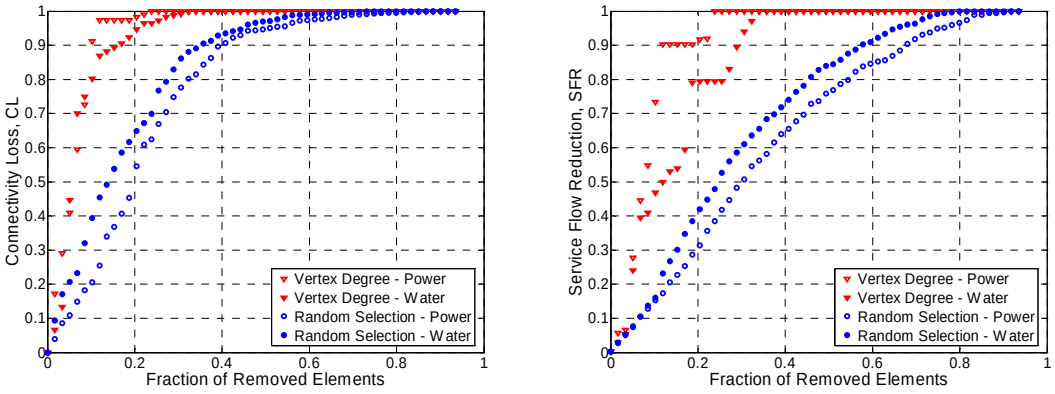


(a) Power grid (b) Water network

Figure 6-5. Service flow reduction, S_{FR}, for a power grid and a water network

Figure 6-6 presents, for each response, the variation in performance as bounded by the random and targeted degree removals. Something additional portrayed by these sets of response measures is that the performance of the networks is expected to be closer to the random selection bound. If vertex transshipment flow captures the optimal flow that traverses each node to meet the constrains of supply and demand, then the response

based on transshipment better reflects the flow process that occurs within real transmission and distribution systems. This is clearly evident for removal fractions less than 30%.

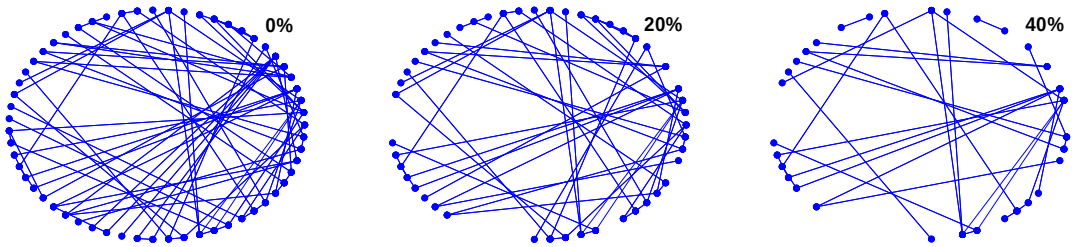


(a) Connectivity loss (b) Service flow reduction

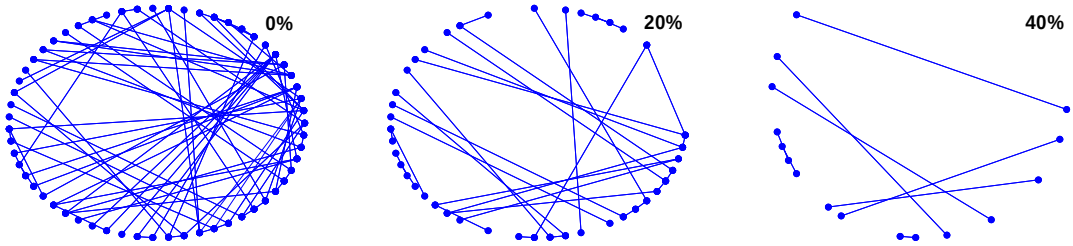
Figure 6-6. Bounded network response: (a) C_L and (b) S_{FR}

One question triggered by the analysis of these global response plots is what does the topology of the disrupted network looks like? The plots clearly show that network functionality is severely impaired by targeted removals and less impacted by random removals. Other removal strategies lie somewhere in between. However, it remains to be understood how the networks fragment. Figure 6-7 presents sequential states of the power network for various fractions of removed elements (i.e., 0%, 20% and 40%), and for the two bounding removal strategies (i.e., random selection and targeted vertex degree). The plots are an abstract representation of the spatial network connectivity. Disconnected nodes are not shown.

It is clear that for significant removal fractions, the strategy based on vertex degree is more effective at partitioning the network. This targeted strategy severely disconnects the network, and fewer remaining elements exist in several isolated islands. For example, at the 40% removal fraction, random removals retain more elements, and they are kept as a larger connected cluster. Vertex degree removals deplete the edges of the system faster and induce a fractured remaining network.



(a) Random vertex removal

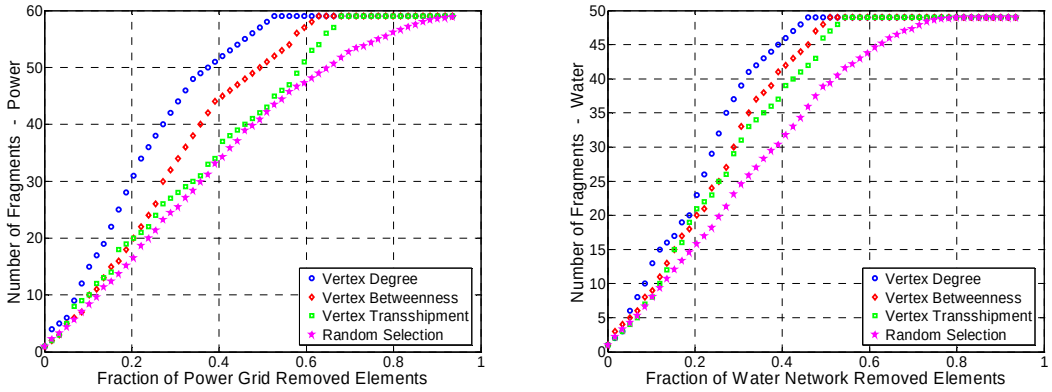


(b) Vertex degree removal

Figure 6-7. Fragmentation patterns for: (a) random and (b) vertex degree removals

These abstract network representations call for a more systematic monitoring of the number of fragments, N_f, in which the network partitions as a result of the removal strategies. Figure 6-8 presents the fragmentation evolution of the networks. The rate of

fragmentation tends to be similar for all removal strategies. Every removed node also reduces the possibilities of its neighbors to remain connected, exacerbating the detrimental effects.

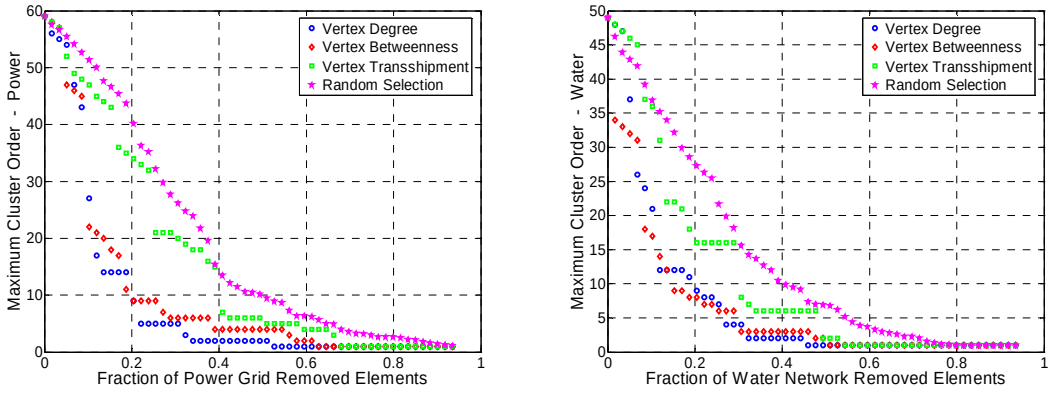


(a) Power grid (b) Water network

Figure 6-8. Network fragmentation for a power grid and a water network

A natural question that follows from the insights on fragmentation patterns is what would be the order, n, of the largest connected cluster? If a disturbance occurs and vertices are removed systematically, the larger the order of the remaining clusters the better chances the network has to largely remain functional. The largest cluster order, n_c, also indicates the potential of the network to restore functionality if it has been lost. Reconnecting small islands to the largest dominant cluster is easier than putting together several small isolated clusters with no clear indication of which is the dominant one. Figure 6-9 presents the order evolution of the largest cluster order. All vertex removal strategies except random selection are effective at breaking down the structure of the networks. However, the strategy based on optimal flow traversal (i.e., transshipment flow) is still closer to the response delineated by random disruptions.

Real systems, despite having some nodes with more connections than others, tend to fracture at locations where nodes are not necessarily special in terms of their vertex degree. Flow transfer in these transmission and distribution networks is made optimal by utilizing as many routes as possible. The fact that the cost of flow traversal is a convex function of the amount of flow calls for diversity in the portfolio of transmission paths to meet flow demands. This makes the network resilient to congestion, and if separation is enforced, the effect of removing highly *loaded* nodes—where load refers to the amount of flow traversing the node—is not as dramatic as if the most *connected* nodes were removed.



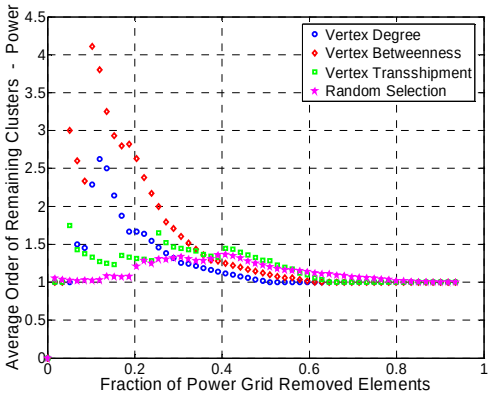
(a) Power grid (b) Water network

Figure 6-9. Maximum cluster order for a power grid and a water network

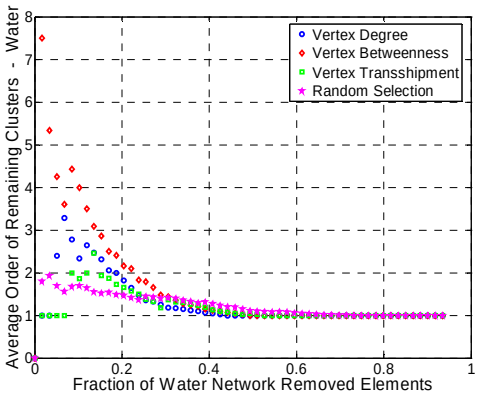
These observations are supported by another metric: one that monitors the average order, $\langle n_r \rangle$, of the clusters without including the largest one. The plots in Figure 6-10 indicate that whenever the network gets disconnected, the average order of the disconnected portions is minimal, $\langle n_r \rangle \leq 2$, for the random and transshipment based

removal strategies. Removing the most connected nodes leaves larger islands. However, removing the nodes with the largest betweenness leaves a significant average order for the remaining clusters. This is interesting because the vertex betweenees removal sequence is finding *bottlenecks*, which when removed can separate the network into similar size clusters. If the removal sequence based on transshipment flow were calculated as a minimization of costs as a linear function of the amount of flow, instead of a convex function, then the sequence would be comparable to that of the vertex betweenness strategy.

It is important to confirm whether or not the response of real systems to systematic removal of nodes is closer to that for a random removal strategy. The average vertex degree reduction is another parameter that reinforces this notion, Figure 6-11. The behavior of this parameter to vertex removal based on transshipment flow which captures real flow distributions is similar to that for random vertex removals. This effect is more pronounced in the power network. The reason goes back to the vertex degree distribution of the networks. Picking and removing *typical* nodes from the water network is more significant than doing so for the power network which has more small-degree nodes due to its positively skewed distribution.

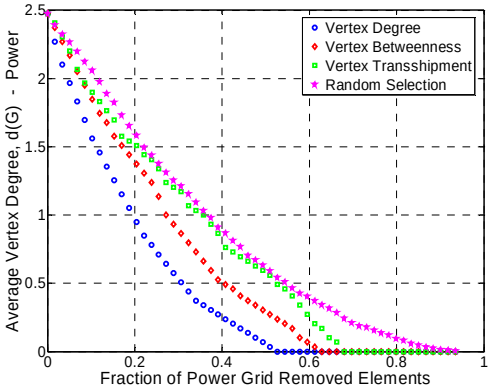


(a) Power grid

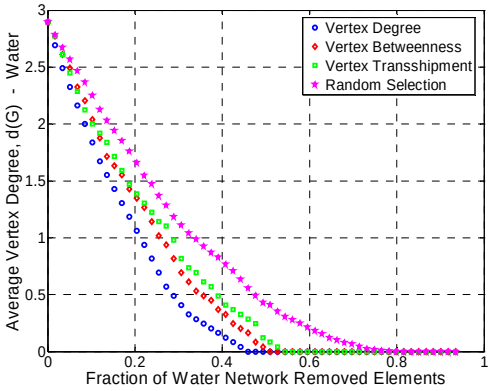


(b) Water network

Figure 6-10. Average order of remaining clusters for a power grid and a water network



(a) Power grid

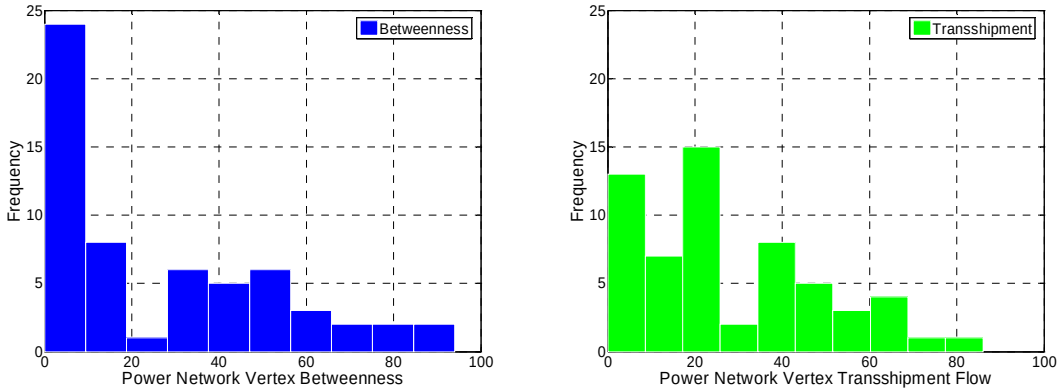


(b) Water network

Figure 6-11. Average vertex degree for a power grid and a water network

All the performance measures and additional topological metrics used to characterize network response to disruptions have a common feature: their monitored responses are bounded by the random and targeted vertex degree removal values. Why does that occur?

To answer this question, it is worth looking at the distribution of the vertex betweenness and the vertex transshipment of the networks at their intact state. Figure 6-3 presents their vertex degree distribution. Figure 6-12 displays the distribution of the vertex betweenness and vertex transshipment flow for the power grid of Shelby County, TN, at its undisturbed state. There are 8 generation units in the generation subset $F_P \subseteq G_P$. The order of the distribution subset $Q_P \subseteq G_P$ is $n_{QP} = 37$. Hence, the maximum possible betweeness is $B_v = 296$. The actual measured B_v ranges from 1 to 100. The vertex transshipment flow, W_v, is a metric proportional to the population size of the County. Each element of the distribution subset Q_P has a power demand associated with it. The power demand at each distribution node is treated as a fraction of the total power demand of the County. Hence, for solving an integer convex program to calculate the optimal flow patterns the smallest fractions are multiplied by 1,000 so that no fractions need to be handled at the demand nodes. The maximum possible transshipment would be $W_v = 1{,}000$. The measured W_v for this power network ranges from 1 to 100.



(a) Betweenness (b) Transshipment

Figure 6-12. Power grid vertex distribution for: (a) betweenness and (b) transshipment

Table 6-1 presents a summary of the main descriptors for these vertex parameters. In essence, the impact on network performance decreases in severity from targeted vertex degree removals, to vertex betweenness, to vertex transshipment, and to random selections. The key to explaining this lies in the distribution of these vertex parameters which, in turn, is the basis of their removal sequences. For obtaining a rank-ordered sequence of vertex removals, the tails of the distribution indicate where to find the most important initial nodes. If the criterion is vertex degree, there are few *special* nodes (i.e., more connected tan other nodes) in the fast decaying thin tail. If they are removed they cause an immediate large impact on network performance, and this is persistently displayed in the response plots. If the criterion is vertex betweenness, then there are more vertices along the tail with high betweenness as indicated by its positive skewness and large COV. These special vertices on the tail of the vertex betweenness distribution, when removed, can cause significant impact on network performance. However, they are more common than in the vertex degree case, thus lessening their detrimental effect if absent from the network. Finally, if the criterion is vertex transshipment, then there are even more vertices along the shorter and thicker tail of the distribution. Its smaller skewness indicates that there is a tendency of having less special nodes in terms of transshipment. Hence, their removal, despite inducing a disruptive effect in network performance, is not as critical as in the other cases.

The tail behavior of the distributions can be further explained using the information captured by the *kurtosis*, k. This parameter measures how prone to outliers a distribution is. It corresponds to the normalized fourth moment of the distributions. The

size of the sample is $n_p = 59$. Data that follows a normal distribution has $k = 3$. The data of the vertex transshipment flow has $k_{Wv} = 2.62$, while for the vertex degree is $k_{d(v)} = 3.74$. These numbers are relative to the expected tail of a normal distribution. This indicates that the W_v is less outlier-prone, hence its thicker and shorter tail, and that $d(v)$ is more prone to outliers, hence its thinner and longer tail. In the case or random selection the probability of picking and removing important nodes is an inversely proportional function of the order of the network. Chance makes more difficult the removal of critical network elements.

Table 6-1. Statistical characterization of power grid vertex distributions

Vertex Property	Fundamental Descriptors				
	Mean μ	**Standard Deviation** σ	**COV**	**Skewness Coefficient** ϕ	**Kurtosis** k
Vertex Degree	2.47	1.07	0.43	0.66	3.74
Vertex Betweenness	25.66	27.17	1.05	0.75	2.45
Vertex Transshipment	27.13	21.57	0.79	0.62	2.62

In spite of the differences in network response induced by the type of disturbance, some disruption mechanisms display similar trends. This observation leads to further investigation of their correlations. Each removal strategy is established by rank-ordering the vertices of the network at the undisrupted state. Then, after removing the most significant element, the rank-ordering is reassessed. This is repeated until all nodes are

removed. The sequence of selected vertices at every stage provides the arrangement for systematic vertex removal. Figure 6-13 displays the correlation between removal sequences of three removal strategies. The vertex betweenness sequence is selected as the reference removal strategy. This sequence shows that its induced disturbances are in general between the disturbances generated by vertex degree removal and vertex transshipment flow. Its correlation coefficient with transshipment is $\rho_{Bv,Wv} = 0.58$, whereas with the recalculated vertex degree is $\rho_{Bv,d(v)} = 0.36$. Its higher correlation with vertex transshipment is expected since both removal mechanisms rely upon flow traversal.

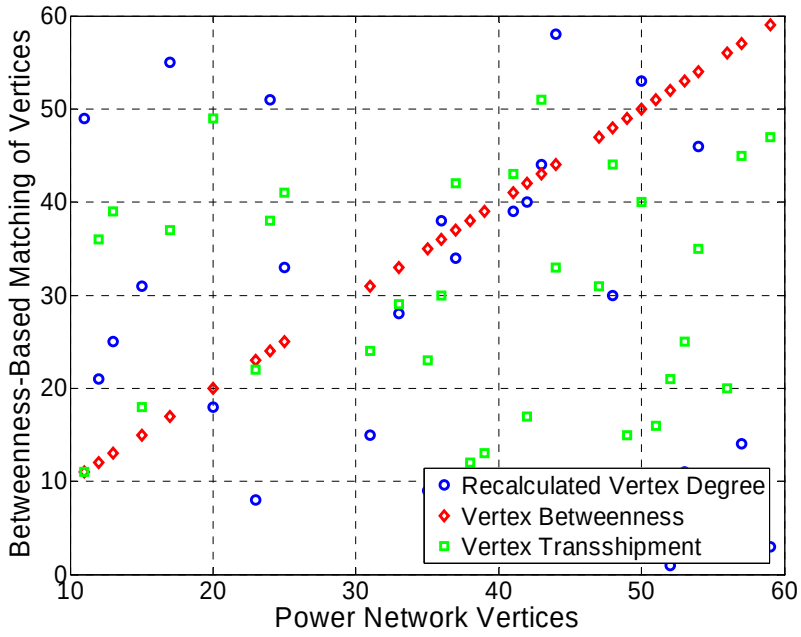


Figure 6-13. Correlation between vertex removal sequences for the power grid

For completeness, the following correlation matrix summarizes the correlation between all removal sequences, plus a removal strategy based on just the initial vertex degree without recalculating it after each vertex deletion:

$$\rho = \begin{bmatrix} 1.00 & 0.22 & 0.06 & -0.11 & -0.03 \\ 0.22 & 1.00 & 0.21 & 0.36 & 0.36 \\ 0.06 & 0.21 & 1.00 & 0.17 & 0.01 \\ -0.11 & 0.36 & 0.17 & 1.00 & 0.58 \\ -0.03 & 0.36 & 0.01 & 0.58 & 1.00 \end{bmatrix} \quad \textbf{(6-1)}$$

where the sequences are: $\rho_{1,1}$ = initial vertex degree, $d(v)_{\text{initial}}$; $\rho_{2,2}$ = recalculated vertex degree $d(v)$, $\rho_{3,3}$ = random selection, $\rho_{4,4}$ = vertex betweenness, B_v; and $\rho_{5,5}$ = vertex transshipment, W_v.

A final check to confirm the variations in network response induced by the different removal strategies is to monitor the reduction of the network size, m. This edge counting gives an indication of the amount of links that are lost after every node removal. Losing edges implies losing the possibility to reroute flow, and therefore making less desirable the measured network properties. The trend shown in Figure 6-14 is unequivocal. Network size is bounded by the random and the vertex degree removal strategies. Even the sequence based on only the information of the network vertex degree at its undisturbed state is more harmful than the sequences based on simultaneous flow and topology. The expected response of the network if a cascading failure occurs—as captured by the transshipment flow—is closer to the response of the network if random elements were removed. Flow distribution and redistribution according to optimal supply/demand constraints causes the network to be more *egalitarian*: no single node becomes dominant or completely essential for performing the intended transmission of flows. This can be generalized to any transmission and distribution network whose flow is indivisible and traverses the system by transfer and not by replication. Hence, similar observations apply to the water network.

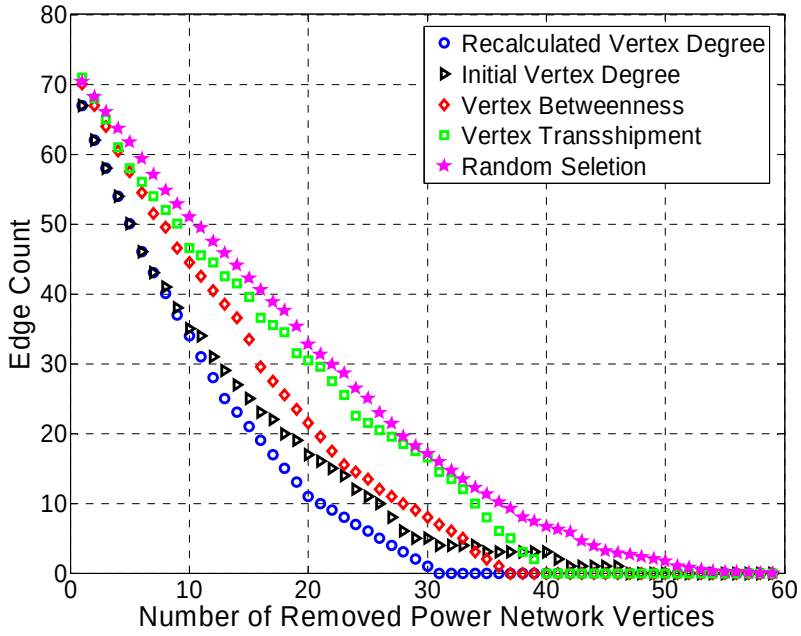


Figure 6-14. Reduction in power network size, m_P, for selected node removal sequences

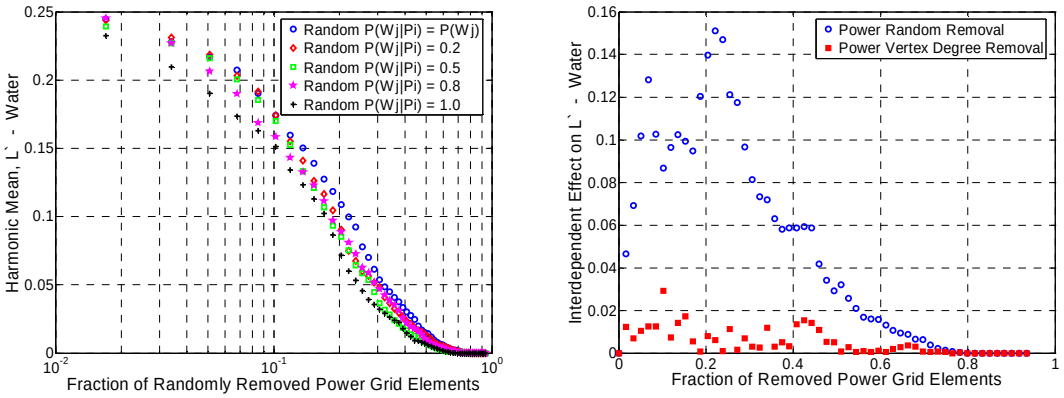
6.1.2 Interdependent networks

For the two example networks (power and water) in Shelby County, TN, introducing interdependencies means $P(W_j|P_i) > P(W_j)$. For the extreme case of $P(W_j|P_i) = 1$, any failure in a power grid element will automatically disconnect the water elements that interact with it, if they exist. Failure in the power grid can occur following any of the disruptive strategies already investigated: vertex degree deletion, vertex betweenness removal, vertex transshipment flow removal, and random selection of vertices for removal. Figure 6-15a presents the response of the water network—measured by its reciprocal harmonic mean, L`—as a function of its strength of coupling with the power grid. Failure in the power grid and in the water network is induced by the random vertex selection. The plot indicates that accounting for interdependent effects results in an

increase of the negative impacts on global network connectivity. The plot enlarges the zone of low fraction of removed elements, < 50%, where most lifelines operate most of the time: within normal or mildly abnormal conditions. Larger removal fractions are induced by less frequent natural hazards such as cyclones or earthquakes. Figure 6-15b presents a quantification of the effect of network coupling on the response and it introduces a new metric: the *interdependent effect, I_e*. This parameter is defined as the absolute difference between the independent and interdependent response. This difference is normalized by the maximum independent response attained for all removal fractions:

$$I_e = \frac{|\text{Interdependent Response} - \text{Independent Response}|}{\max(\text{Independent Response})} \quad \textbf{(6-2)}$$

The plot includes the effect of power grid failure on water network response. This effect on the response is monitored for the power grid undergoing disruptions according to the two bounding removal strategies (i.e., random and vertex degree removals). Interdependencies are critical in network response for removal fractions < 40%. Their magnitude is also substantial. At its peak, the interdependent effect accounts for an additional reduction of 16% in network functionality with respect to its intact state.



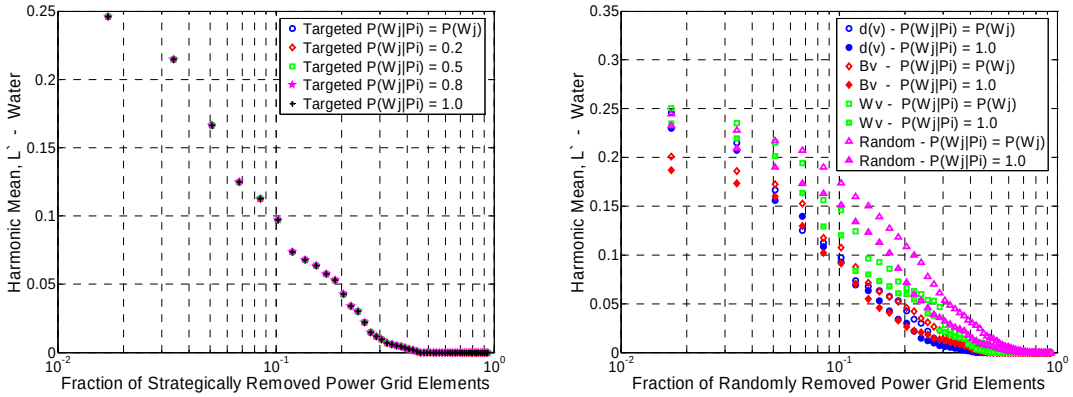
(a) (b)

Figure 6-15. Interdependent network response: (a) *L'*, and (b) I_e on *L'*

An interesting outcome of the interdependent effect plot comes from the targeted removals in the power grid. Picking the most connected nodes of this system appears to cause no significant interdependent effects on the water network. If the power network is being severely disrupted, then why is the water network indifferent? The reason is simple: the most connected nodes of the power grid are not necessarily the nodes that facilitate the interaction with other systems. At the interacting interface, the water network knows of power nodes which are not the specially connected ones.

Since the impact that a particular removal strategy provides interesting clues about network behavior, Figure 6-16 presents additional cases. Figure 6-16a displays the response of the water network when both the power and the water network are subjected to targeted removal of their most connected nodes. The interdependent effects now vanish completely. No strength of coupling is influential to the response. This is because both networks are prematurely degraded, leaving their remaining interface elements unable to add any detrimental effects. Figure 6-16b shows the response of the water

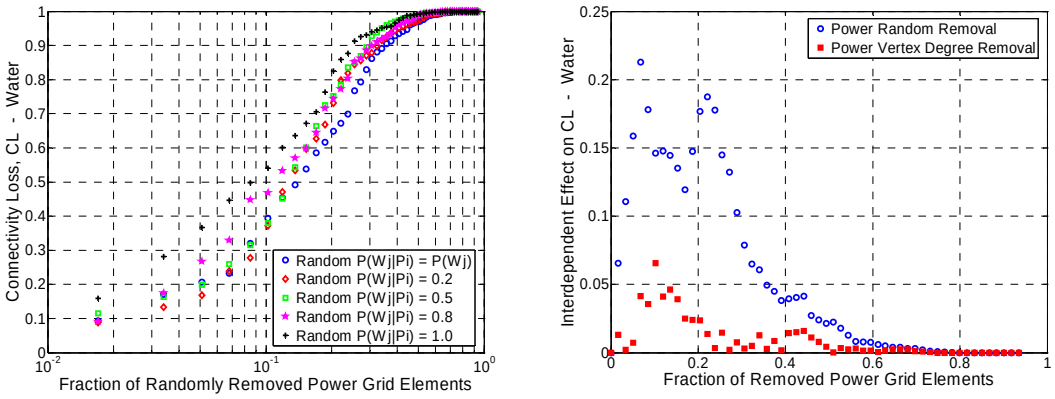
network when it is subjected to each of the vertex removal strategies while the power grid is undergoing only random removals. The strength of the coupling, as characterized by the conditional probabilities of failure $P(W_j|P_i)$, is monitored for its limiting values (i.e., $P(W_j|P_i) = P(W_j)$ and $P(W_j|P_i) = 1.0$). It confirms the trends exposed by the independent network analysis. The interdependent response of the networks is also bounded by the random and the targeted vertex degree, $d(v)$, removal strategies. The effect of removal strategies based on simultaneous topology and flow (i.e., vertex betweenness, B_v, and vertex transshipment flow, W_v) lies between the bounding strategies.



(a) (b)

Figure 6-16. Interdependent *L'* response for: (a) coordinated attacks, and (b) all type of removals

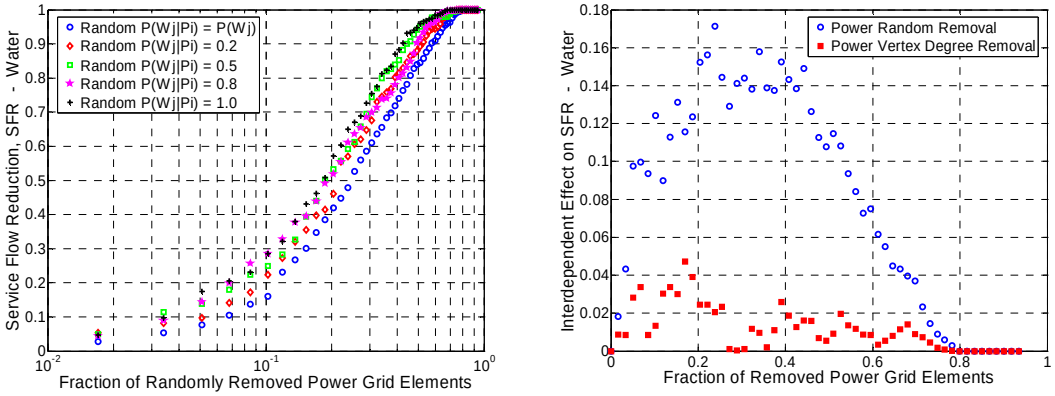
Figure 6-17 presents the interdependent response of the water network in terms of its connectivity loss, C_L. Again, the strength of coupling accelerates the loss of connectivity between distribution and generation nodes. The effect of the interdependencies on the response is sizable for removal fractions < 40%. The magnitude of the interdependent effect at its peak amounts for 22% of the independent response.



(a) (b)

Figure 6-17. Interdependent network response for: (a) C_L and (b) I_e on C_L

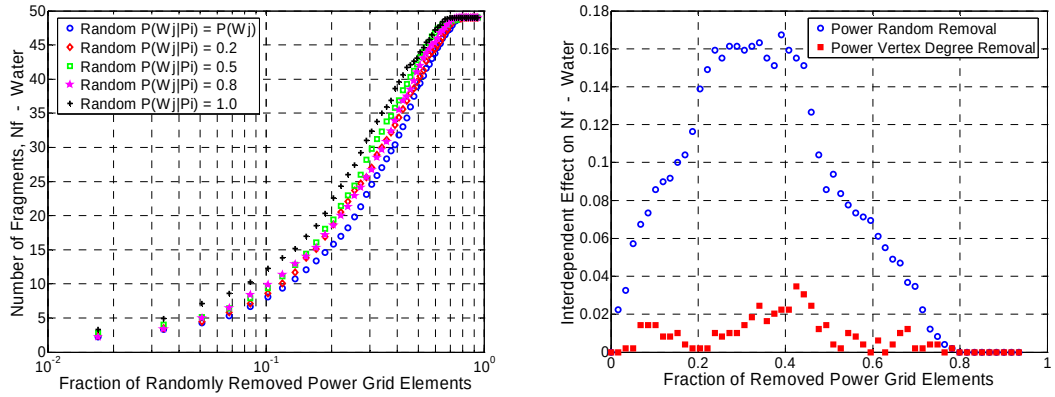
Similar results are observed for the service flow reduction, S_{FR}. Figure 6-18 shows the response of the networks with various strengths of coupling, and its interdependent effects. It is noticeable that the interdependencies are significant for a larger portion of the removal fraction range. Since the requirements for optimal flow distribution induce a widespread utilization of all network elements, the interdependencies are significant almost up to the point in which the network is going to completely break down.



(a) (b)

Figure 6-18. Interdependent network response: (a) S_{FR} and (b) I_e on S_{FR}

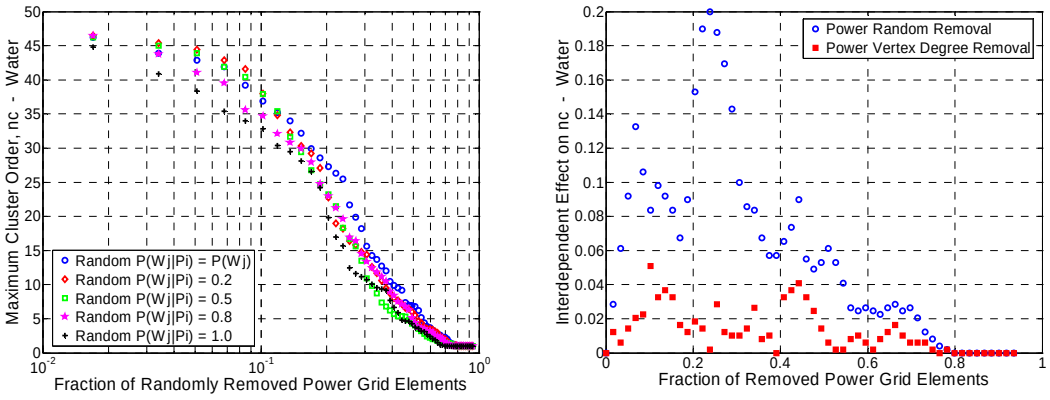
The number of fragments in which the network breaks displays the same trends of the S_{FR} response metric (Figure 6-19). The interdependent effects remain active well above mild removal fractions. This is because as the network partitions into islands due to vertex removal, any node becomes more likely to disconnect. Additional disturbances coming from the interaction lead these nodes to eventual failure.



(a) (b)

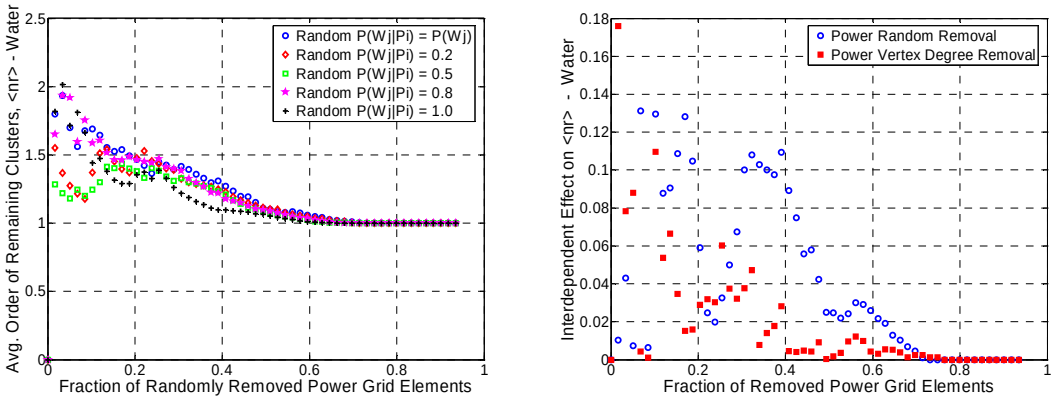
Figure 6-19. Interdependent network response: (a) N_f and (b) I_e on N_f

To complement the information provided by the fragmentation patterns of the networks, Figure 6-20 and Figure 6-21 present the reductions of the network order, n, and the average order of the fragments without counting the largest one. Interdependencies reduce the order of the largest cluster. The interdependent effect, I_e, is significant for a substantial portion of the fragmentation removal range. The two plots suggest the following: for small removal fraction the preferred failure mode of interdependent networks consists of multiple fragmentations, where the distribution of the order of the clusters is highly skewed. This means that most of the fragments are small $n < 2$, with the exception of one that tries to keep a maximum size ~ n_W—the original water network order. The implications of this observation can have a large impact in devising consequence minimization actions that take advantage of this well behaved largest cluster. Speed of recovery is a function of the ability to reconnect the entire network. The backbone for this reconnection is the maximum order cluster.



(a) (b)

Figure 6-20. Interdependent network response: (a) n_c and (b) I_e on n_c



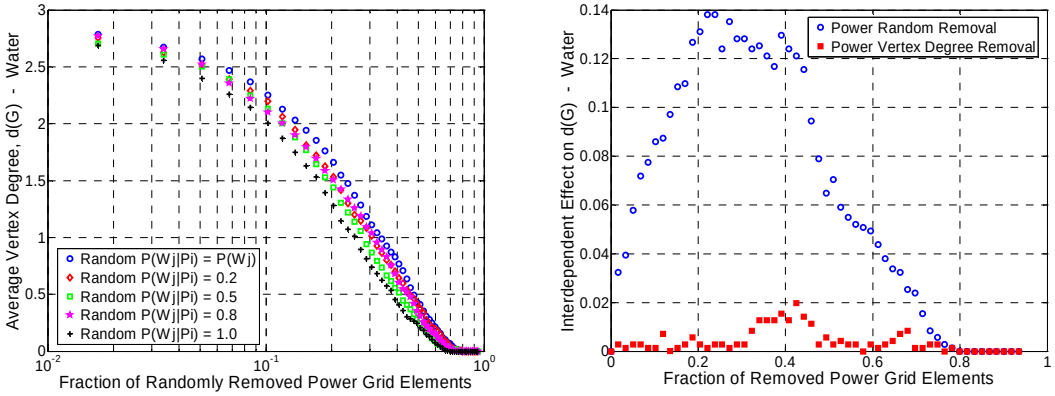
(a) (b)

Figure 6-21. Interdependent network response: (a) $<n_r>$ and (b) I_e on $<n_r>$

A final response metric for monitoring interdependent response is the average vertex degree of the entire network, $d(G)$. Figure 6-22 shows that the overall vertex degree diminishes faster in the presence of interdependencies. These effects are felt for the entire removal fraction range. Every node counts and each removal preconditions more nodes to be depleted of its connections. Additional disturbances induced by interacting systems simply exacerbate the incident edge count decrement.

All monitored network response metrics, either those based upon topology or simultaneous topology and flow, show a common trend. The strength of the interdependencies accelerates non-desirable properties in the networks. These effects are bounded as expected by the independent and interdependent states (i.e., limiting values of the conditional probabilities of failure $P(W_j|P_i)$). The effect of the interconnection is more significant for removal fractions < 50%. However, some metrics that are closer to capturing physical processes within the networks, such as S_{FR}, show significant

interdependent effects up to removal fractions of 80%. The magnitude of the interdependency effect at their peaks was approximately 20% the maximum independent response. This is a non negligible percentage, and it occurs where the fraction of removals is low < 20%. This is precisely the range in which the networks operate, where common faults induce element removals which are small with respect to the network order *n*.



(a) (b)

Figure 6-22. Interdependent network response: (a) $d(v)$ and (b) I_e on $d(v)$

Finally, it is consistently shown that the interdependent effect, I_e, for targeted disruption of one of the networks has little effect on the dependent network. The claim is that the critical nodes for performance and structure of one network are not necessarily the same nodes in charge of transferring the flow at the interface between distinct networks.

6.2 LARGE SCALE PERTURBATION OF NETWORKS

Rather than systematically removing one vertex at a time according to certain criteria to assess node importance this section deals with simultaneous disruptions over large portions of the networks. Earthquakes provide the hazard context. Every element of a networked system experiences different seismic demands depending on its geographical location and its intrinsic dynamic characteristics.

6.2.1 Seismic vulnerability of network elements

Each element of a networked infrastructure (i.e., a system) has an inherent vulnerability to seismic hazards. These systems are in general composed of several subsystems. For instance, a typical element of the power grid, such as an electric substation, contains several subsystems, parts and units, such as transformers, disconnect switches, circuit breakers, regulators, control equipment, and lightning arrestors. Failure of any or several of these subsystems compromises the functionality of the entire substation. Fault tree analyses, numerical simulations, and log-normal fits are tools used by researchers to estimate and describe the conditional probability of failure of network elements to seismic hazards. These probabilities enable development of system fragility curves. Similar system reliability approaches are followed to generate fragility curves for other network elements such as generation units and different voltage electric substations (i.e., high voltage > 350kV; medium voltage, 350kV > voltage > 150kV; and low voltage < 150kV). Vulnerability of systems for other networks is established in analogous fashion. Water distribution network elements include large (> 10 million gallons per day, mgd) and medium

(< 10mgd) pumping plants, elevated storage tanks (~ 2mgd), and brittle or ductile delivery pipelines. Generic estimates of network element fragility relationships are available from various sources, for instance HAZUS-MH (Federal Emergency Management Agency, 2005). Most of the data for calibration of these damage functions comes from studies of network elements on the west coast of the United States where the seismic activity is larger than in the central/eastern part of the country. Figure 6-23 displays the probability of exceeding either an extensive or complete damage limit states as a function of peak ground acceleration (PGA) for the major elements of a power grid. Figure 6-24 displays the vulnerability for the essential elements of a potable water network. The *extensive damage* limit state implies damage beyond short term repair, leaving the network systems under consideration as completely nonfunctional. The *complete damage* limit state implies collapse of structures and failure of equipment.

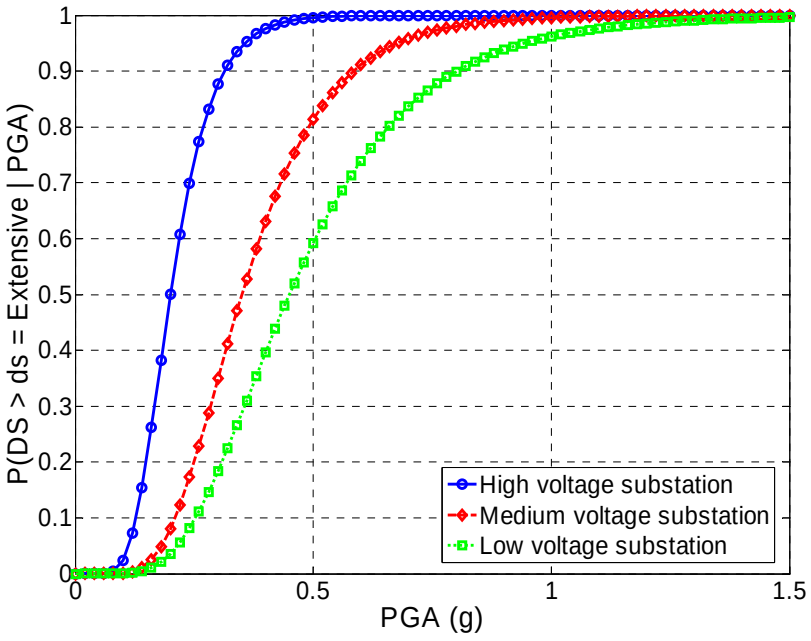


Figure 6-23. Fragility curves for power grid elements exceeding extensive damage

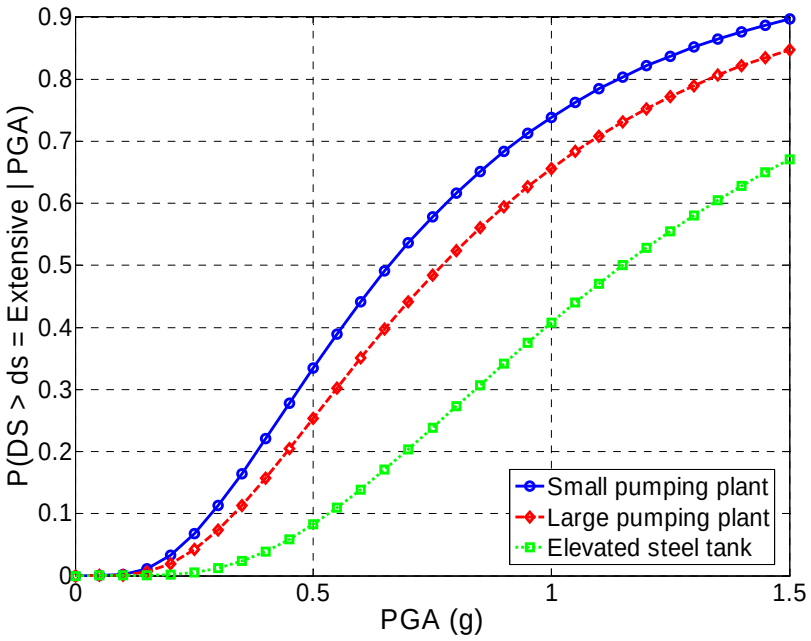


Figure 6-24. Fragility curves for water network elements exceeding extensive damage

Damage functions for buried pipelines without soil susceptibility to liquefaction, have empirical evidence that they follow power laws (i.e., linear trend in log – log scales). These functions usually estimate the rate of repair or expected number of repairs per unit of length as a function of hazard intensity. The equation below shows a sample relationship in terms of peak ground velocity (PGV). This empirical relation was developed by using post-disaster data from past United States and Mexican earthquakes (O'Rourke and Ayala, 1993):

$$R_{rate} = \frac{1}{10{,}000}(\mathrm{PGV})^{2.25} \qquad \textbf{(6-3)}$$

where the repair rate, R_{rate}, corresponds to the number of repairs per kilometer of brittle pipelines, and PGV is in cm/s. Repair rates are caused by two major damage mechanisms: breaks and leaks. In general, 15 - 20 % of those damage mechanisms result in breaks or ruptures, while the rest represent leaks (Hwang et al., 1998). The diameter of the pipelines ranges from $\phi = 0.15$m to $\phi = 1.22$m in order to include transmission and distribution links. This R_{rate} parameter is useful to characterize the probability of having pipeline ruptures because it allows estimation of the mean occurrence rate of breaks. This characterization is done by introducing a spatial Poisson process to define the probability that the number of pipe breaks B_r equals r within a given pipeline segment of length L:

$$P(B_r = r) = \frac{(R_{rate}{}^{*}L)^r}{L!} e^{-R_{rate}{}^{*}L} \quad \textbf{(6-4)}$$

In this study, it is assumed that the occurrence of at least one break will impair the functionality of a pipeline segment. Therefore, the probability of pipeline break occurrence simplifies to:

$$P(B_r > 0) = 1 - P(B_r = 0) = 1 - e^{-R_{rate}{}^{*}L} \quad \textbf{(6-5)}$$

This equation when applied to all pipelines of a junction permits calculation of the failure of pipeline *intersections* given a PGV level of seismic hazard. These intersection nodes are fundamental part of the topological characteristics of water networks according to the graph theory approach of this study. In essence these junctions lump the incident pipeline fragilities. Non-functionality or failure, F_v, of a junction, v, is reached when all

pipeline segments incident to v have at least one break. This is analogous to the probability of failure as prescribed by the reliability of parallel systems:

$$P_f = P(F_v) = P(F_{v_1} \cap F_{v_2} \cap \ldots \cap F_{v_{d(v)}}) \quad \textbf{(6-6)}$$

where the vertex degree, $d(v)$, corresponds to the maximum number of incident segments of the intersection under consideration. This relationship, along with the fragility curves of Figure 6-23 and Figure 6-24, fully characterize the vulnerability of fundamental constitutive subsystems and parts of power and water networks. Other network elements exist, but the ones already introduced (e.g., electrical substations, pipelines, tanks, and pumps) are sufficient to characterize the simplified power and water networks of Shelby County, TN.

Figure 6-25 presents samples of *generic* pipeline junction fragility curves for a nonfunctional damage state *DS* (i.e., at least one break per pipeline segment). The curves are shown for a range of typical topological parameters of main pipes of water distribution systems (e.g., average pipe length and average vertex degree). The water network used in this study has a vertex degree $d(v)$ that ranges from 1 to 5, with an average of $d(G) = 3$. Also, pipeline segment lengths, L_{vj}, from vertex v to its j^{th} immediate neighbors, vary from 1 km to 15 km, with an average $\langle L \rangle = 6.9$ km.

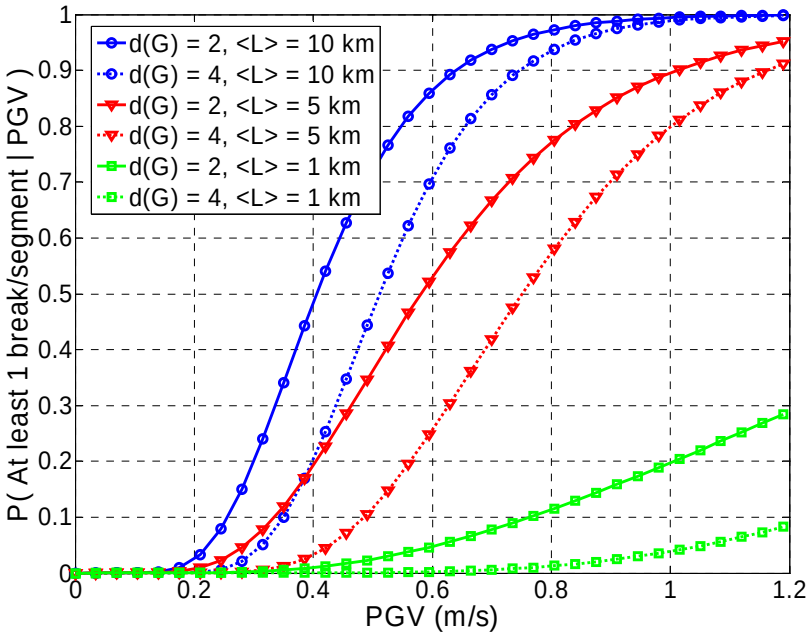


Figure 6-25. Generic fragility curves for brittle pipeline junctions as a function of PGV, average vertex degree $d(G)$, and average pipeline segment lengths $<L>$

6.2.2 Seismic hazard and independent network vulnerability

In order to develop fragility curves for interdependent networked systems, it is necessary to evaluate the network performance under several levels of seismic hazard. Since the constitutive network elements are distributed over a large geographical area, the levels of seismic hazard to which they are subjected are different depending on their location. Therefore, it is necessary to describe the temporal and spatial characteristics of the disturbance (i.e., hazard). In the present study, PGA and PGV contours for Shelby County are established for several earthquake return periods using HAZUS-MH. This hazard definition follows the probabilistic methodology proposed by the United States Geological Survey (USGS), (Frankel et al., 2002). This methodology for the Central/Eastern United States uses parameters such as magnitude, distance and return

period to determine the contribution to the hazard from different seismic sources. These sources include historical seismicity, large background sources, and specific fault sources. Figure 6-26 and Figure 6-27 present a sample of the seismic hazard in PGA and PGV for Shelby County, TN. The maps display seismic hazard contours for a hazard rate consistent with a 10% probability of exceedance in 50 years. The PGA and PGV contours are matched with the boundaries of the census tracts, and the main components of the water and power networks are displayed without transmission or distribution links for clarity.

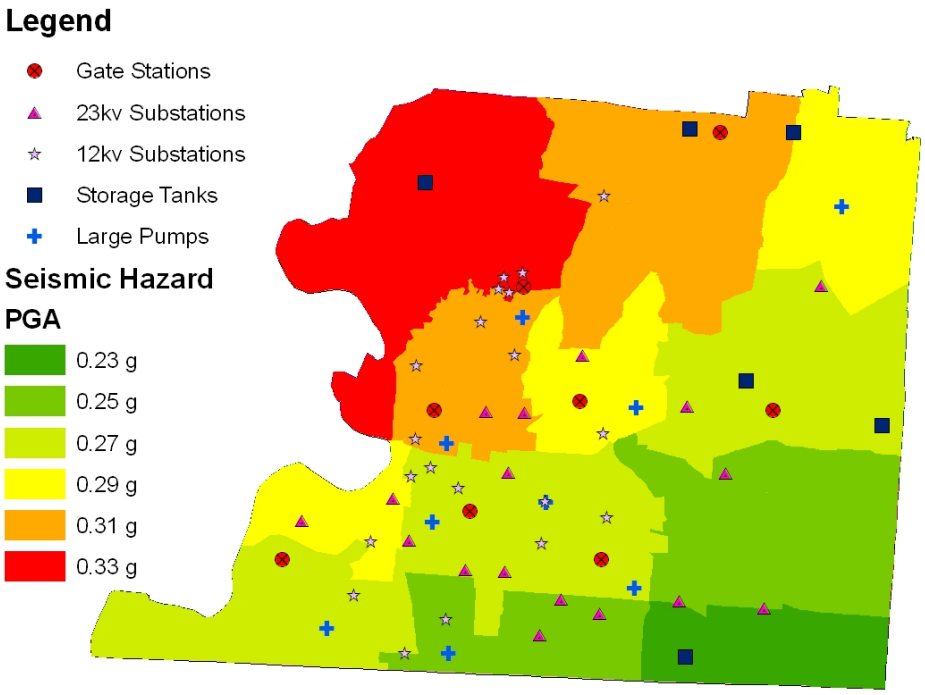


Figure 6-26. Probabilistic seismic hazard for Shelby County, TN, in PGA

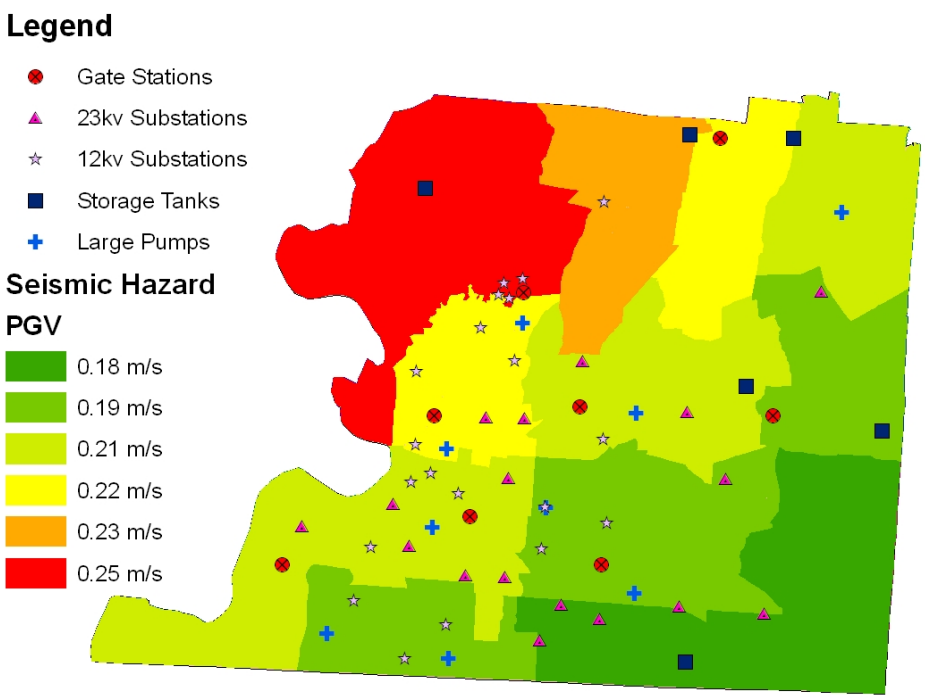


Figure 6-27. Probabilistic seismic hazard for Shelby County, TN, in PGV

The geographical distribution of PGAs and PGVs as the hazard propagates to the southeast portions of the county decreases relative to the maximum observed values in the northwestern part of the county. This trend is similar for all return periods under consideration (i.e., 100, 250, 500, 750, 1,000, 1,500, 2,000, and 2,500 years). The expected PGA and PGV values are recorded for each seismic hazard return period at the location of each network element. Since the probabilities of each network element to exceed extensive or complete damage are conditioned on these PGA or PGV values, a Monte Carlo simulation routine is implemented to determine whether a particular network element fails or not. Each element has a probability of failure associated with its location and dynamic characteristics. This probability of failure is compared with a uniform random variable $X \sim U(0, 1)$. If the value of X is larger than the probability of

failure the element, then the element is removed. This is done for all elements at every hazard level and repeated 10,000 times to determine the distribution of the response of the networks. The performance of the network after each trial is measured by the connectivity loss, C_L. After the trials are completed for a particular hazard level, the proportion of all trials that exceed certain network damage state, DS (e.g., $DS > 20\%\ C_L$, $DS > 50\%\ C_L$, or $DS > 80\%\ C_L$) is the limiting probability of network failure given a seismic intensity value PGA or PGV. This is done for various levels of PGA and PGV.

The response data from the simulations leads to the development of fragility curves in terms of the C_L response metric. This parameter is chosen because of its compromise between computational complexity and ability to capture network flow distribution. The worst-case complexity for C_L and efficiency, E, is $O(n^2m)$, while the worst-case complexity for the more desirable S_{FR} response is $O(n^2m^2\log U)$. Also C_L by rudimentarily combining topology and flow is more desirable than the topology-based network efficiency, E.

Figure 6-28 presents a revealing plot for the complement cumulative distribution functions of the C_L. This figure shows that the CDFs for the water C_L display different behavior depending on the hazard level (i.e., approximately 10%, 5% and 2% probability of exceedance in 50 years). The behavior varies according to a power law for low PGA hazard levels, to a linear relation for medium PGA levels, to a logarithmic function for high PGA levels. The effect of interdependencies induces an analogous behavior. The curve for the water network at low seismic hazard, but with 100% interdependence with the power grid or $P(W_j|P_i) = 1.0$, highlights the potential for inducing extra fragility. The CDF for the power C_L at low seismic hazard levels is superimposed in the plot. It

strengthens the notion that both network topology and network element vulnerability have a significant role on the performance of the system.

The power grid proves to be highly vulnerable, despite being the backbone of modern productivity. One of the reasons is that electrical substations which are composed of numerous vulnerable subsystems posses a steep fragility relationship. Also, the order of its generation subset $F_P \subseteq G_P$ is smaller, and its sparseness is higher as compared to the water network. Power demand and generation have experienced a significant growth in the last 25 years, whereas, transmission lines (i.e., edges), growing at slower rates, have been forced to be used beyond their original design capacities. These factors have the effect of reducing the reliability of the system, and of generating the potential for massive blackouts due to transmission overload and subsequent cascading failures (Hauer and Dagle, 1999; United States Department of Energy, 2000).

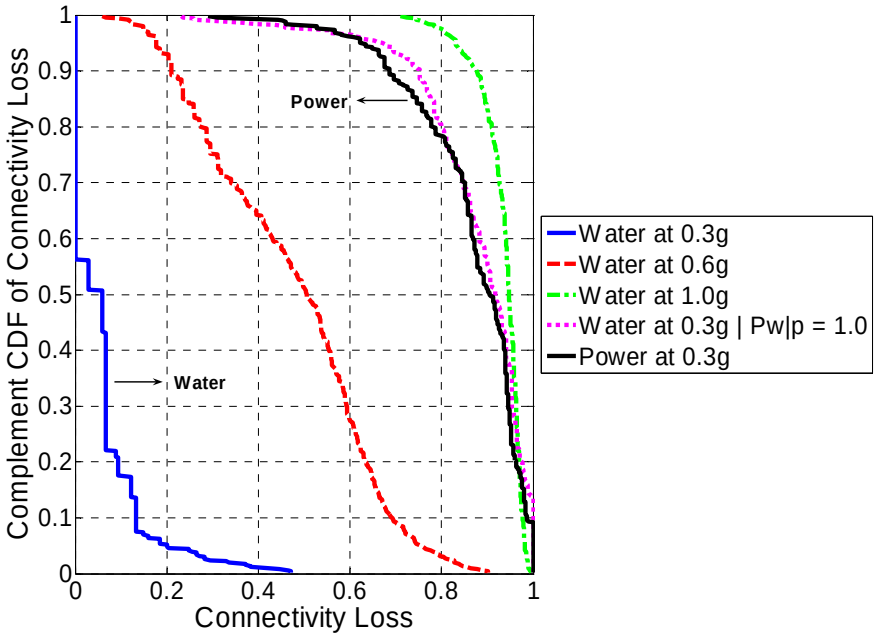


Figure 6-28. Complement of interdependent CDFs for the network connectivity loss, C_L

The distributions of the C_L for several hazard levels provide sufficient information to establish network performance levels. Fragility curves for the networks can be established fitting the probabilities of exceeding specific limit states to a lognormal distribution. Connectivity loss levels C_L of 20%, 50% and 80% represent three limiting states to measure the ability of the network to function properly. More precisely, they quantify, for different performance levels, the likelihood of the distribution nodes to decrease their capacity to be connected to generation nodes given a particular seismic hazard. Figure 6-29 presents the fragility curves for the potable water infrastructure systems as a function of the maximum expected PGA in the area and expressed in term of their connectivity loss, C_L.

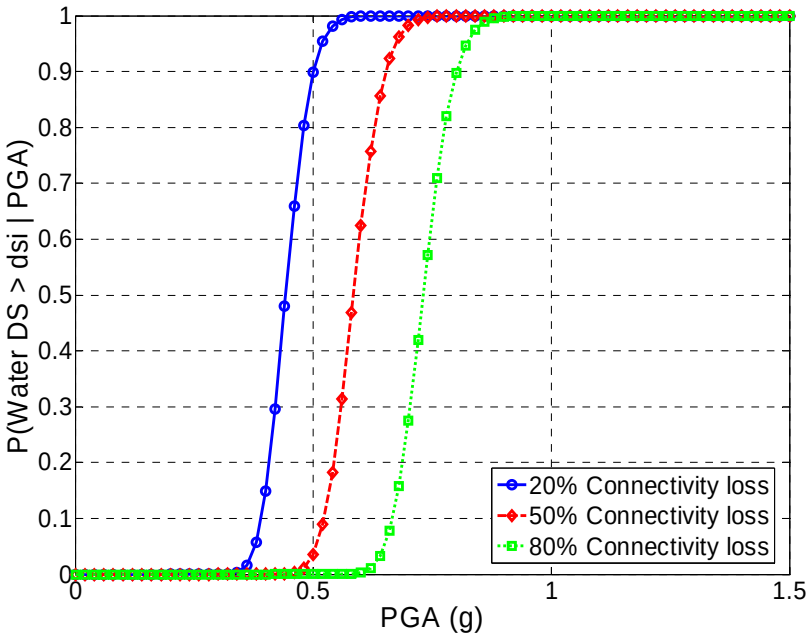


Figure 6-29. Fragility curves for the water infrastructure system as a function of C_L

Figure 6-30 displays the relationships for the electric power grid. The fragility of the power grid, and to some extent the fragility of the water network, reflect the response of a system that undergoes a phase transition: either functions or not. Under disturbances, these transmission and distribution systems try to maintain their global functionality by maximizing the size of the remaining connected portions of the network. As elements fail, this global functionality becomes more and more dependent on fewer links. Once those few links are gone, the network suddenly fragments. This phenomenon is also observed in the growth of abstract random graphs. Figure D-1 in the appendices shows the phase transition in the number of connected of nodes as a function of the total number of edges.

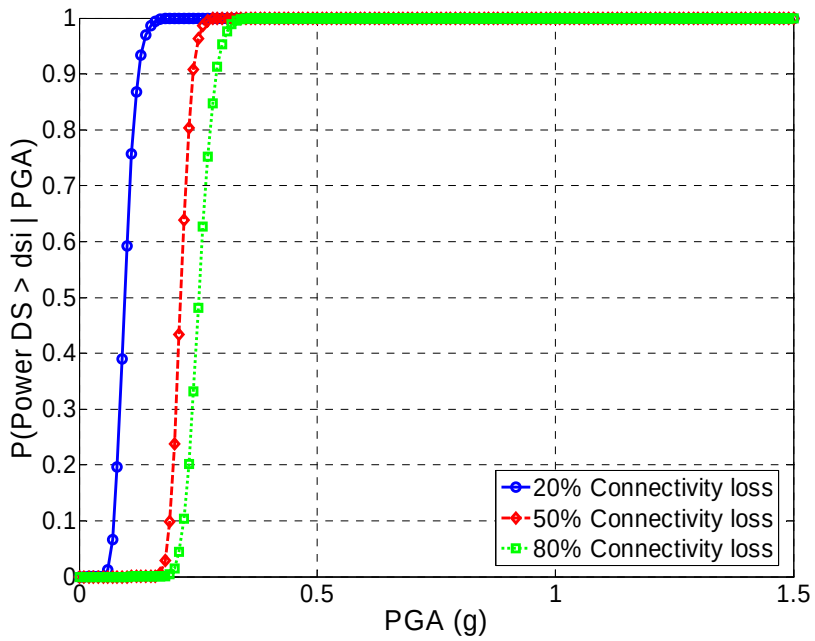


Figure 6-30. Fragility curves for the power infrastructure system as a function of C_L

The next step is to construct system fragility curves after taking into account the effect of network interdependencies. This is done by performing again the Monte Carlo simulation routines of element survival. However, this time for every network element, the algorithm not only compares the probability of failure with a uniform random variable to decide if the element fails or not, but it also compares the strength of the coupling among interacting elements with another random variable. This new uniform random variable $Y \sim U(0, 1)$ is used to determine if the nodes need to take into account the effect of the nodes that condition their performance (i.e., when Y is less than the strength of coupling). The strength of the coupling, denoted by the value $p_{Wj|Pi}$, corresponds to the conditional probabilities of failure of the interdependent adjacency matrix $A_{W|P}$. The following figures introduce a set of *interdependent* fragility curves. They display the effect that seismic-induced failures on the power grid have on the water distribution network. Four different values of the strength of coupling, that is, the magnitude of the interdependence, are considered. These values range from independence where $P(W_{j|}P_i) = P(W_j)$ to complete interdependence where $P(W_{j|}P_i) = 1.0$. Figure 6-31 shows the interdependent effect for a connectivity loss, $C_L = 20\%$. Other performance levels are presented in Figure 6-32 for $C_L = 50\%$, and in Figure 6-33 for $C_L = 80\%$.

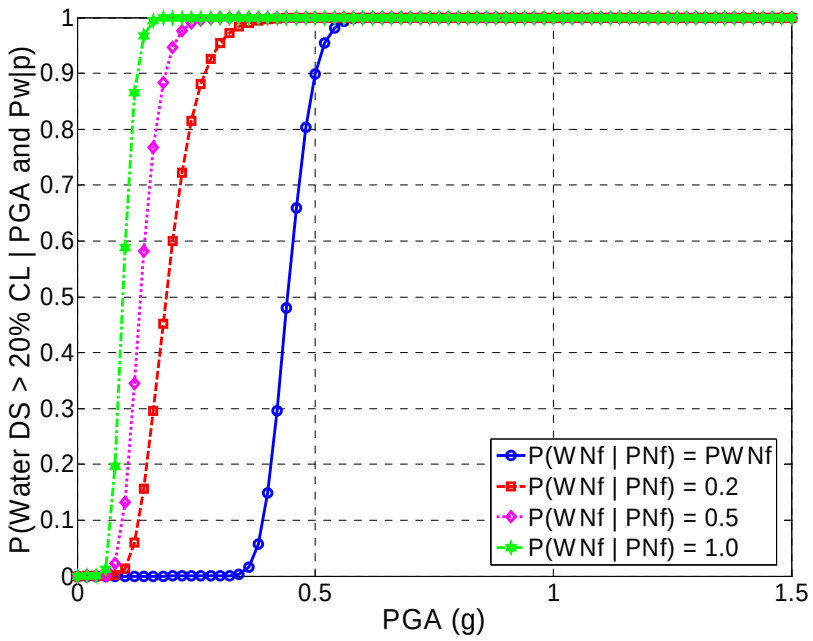


Figure 6-31. Interdependent fragility curves for connectivity loss C_L = 20%

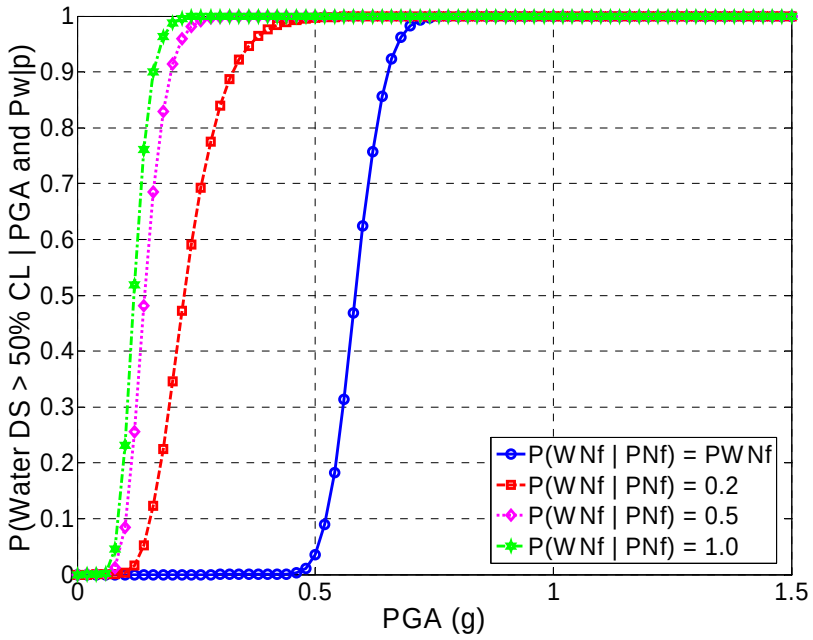


Figure 6-32. Interdependent fragility curves for connectivity loss C_L = 50%

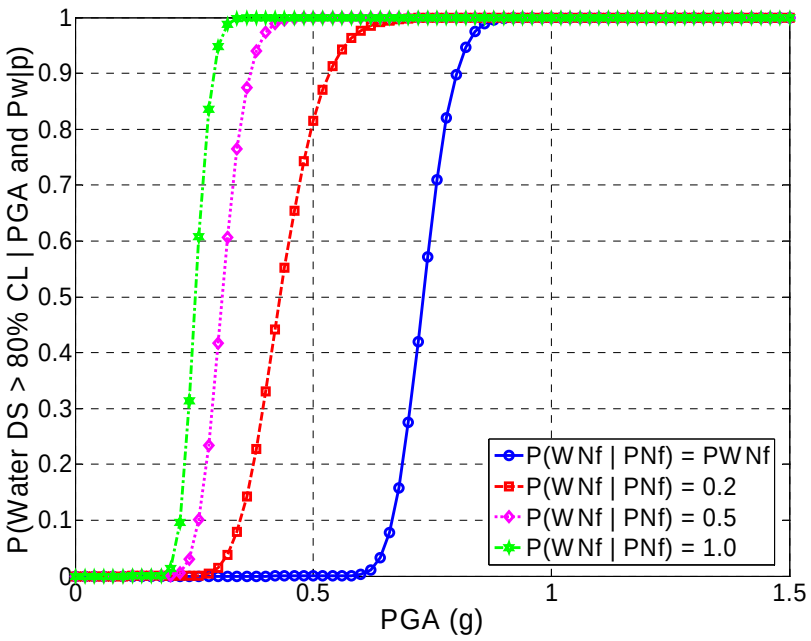


Figure 6-33. Interdependent fragility curves for connectivity loss C_L = 80%

These plots clearly show the increase in system fragility as the strength of coupling grows larger. This effect is amplified as the monitored network performance level, C_L, increases. The median value of the fragility distributions decreases approximately 60% with respect to the median of the independent system at C_L = 20%. The decrements in the median then become 65% and 70% with respect to the independent case for C_L = 50% and C_L = 80%, respectively. The maximum interdependent effect in all performance levels is an astonishing I_e = 100%. Accounting for interdependencies increases the complexity of the systems. This adds new combinations to the possible failure modes, inducing a substantial increase to the overall infrastructure fragility.

A peculiar aspect of these plots is that the marginal contribution of the value of network interconnectedness, $p_{Wj|Pi}$, is relatively small once the interdependencies exist. In other words, the leap from the independent state (i.e., $P(W_j|P_i) = P(W_j)$) to a small

strength of coupling (e.g., $P(W_j|P_i) = 0.2$) is larger than from a small strength of coupling to complete interdependence (i.e., $P(W_j|P_i) = 1.0$). This happens for all the monitored performance levels, C_L. The explanation lies in the insistence that the transmission and distribution systems maintain functionality via widespread utilization of its elements. After a sizeable perturbation, the remaining connected portions of the network tend to remain in a single cluster. This cluster becomes the backbone of the network, but at the same time it is held together by lone cut vertices—nodes whose removal separates the network into isolated components. Introducing any additional perturbation, such as the one induced by interacting with other systems, results in an increased likelihood of losing those nodes that maintain the network functional. Their absence can suddenly fragment the network and increase the overall infrastructure fragility.

6.3 SUMMARY

Analysis of the response of networked systems requires monitoring of their performance parameters as a function of the intensity of disruptive effects. These disturbances are either of internal or external nature. Four strategies for systematic vertex removal are considered: random removal, vertex degree removal, vertex betweenness removal, and vertex transshipment flow removal. These strategies range from simply picking any vertex according to a uniform distribution to a relying upon the actual role of each node to receive and transfer flow under optimality, so that supply and demand constraints are satisfied.

Each of the example networks of Shelby County, TN, networks (i.e., electric power, and potable water) is subjected to systematic disruption of their vertices. Both

networks exhibit similar response as the fraction of removed elements increases. However, the water network displays better performance than the power grid for targeted removals. The opposite occurs for random removals. The reason for that type of response lies in the vertex degree distribution. The more egalitarian degree distribution of the water network (i.e., Poisson with small variation) induces the observed behavior.

Another interesting observation is that all network response parameters are bounded by the evolution of the response to random removals and vertex degree removals. Additional topological metrics that monitor the physical breakdown of the networks also display the same bounds. The response parameters that combine topology and flow are within the stated bounds. However, the service flow reduction parameter, S_{FR}, introduced in this work to capture real flow transfer within networks, has a tendency to remain close to the response to random vertex selection. This indicates that optimal flow requires the utilization of a widespread set of paths to accomplish its distribution tasks. Hence, no specific node is essential for providing the intended function, but rather, every node has a similar share of responsibility, mimicking a random selection process.

The fact that all network response parameters are bounded by the random and the targeted vertex degree removals calls for a deeper investigation. The analysis of the vertex property distributions contains clues to solve the problem. The shape of the tails indicates where to find the special nodes if a systematic removal is planned. Vertex degree removals are more damaging because they result in a few nodes with more connections than others which, when removed, can impair network functionality at a faster rate. Random removal, on the other hand, has less chance of picking these few special nodes. Main descriptors of the vertex property distributions confirm these

observations. Skewness and kurtosis of the sample data capture the tail behavior that leads to the rank-ordering of vertices for systematic removal.

The response of the networks accounting for their interdependencies is also investigated. Each performance metric is examined for several strengths of coupling between networks. In general, the strength of the interdependencies accelerates non-desirable properties of the networks. All network response metrics are bounded by the independent and the fully interdependent strengths of coupling.

Interdependencies are more pronounced when the power and the water network are subjected to systematic random removal. This work introduces a parameter to capture the effect of interdependencies with respect to the independent response. The interdependence effect, I_e, shows that networks have their interacting peak at removal fractions less than 40%. This marks the range where most networks operate under normal and mildly abnormal conditions. Removal fractions beyond 40% are reserved for catastrophic events such as cyclones or earthquakes. When the removal in the power network is strategic, that is, it is based upon vertex degree, rather than random, the interdependent effect decreases. This defies the expected following outcome: the more a system is initially disturbed, the more the detrimental effects are transferred to its interacting systems. This does not happen all the time. The claim is that the critical nodes for performance and integrity of one network are not necessarily the same nodes in charge of transferring the flow at the interface between distinct networks.

The response of the networks is also investigated for seismic hazards. In this case their performance is characterized by the connectivity loss, C_L. This response parameter shows a distribution that follows different regimes depending on the levels of the seismic

hazard. It goes from power law, to linear, to logarithmic functional forms, as the intensity measure increases. The fragility of the infrastructures as a function of the fragility of its constituent subsystems tends to follow a step function. It resembles a phase transition where there is a critical value below which failure does not occur, but above which it does.

The effect of the interdependencies is accounted by explicitly checking the status of each network element for seismic failure and for interacting effects. The connectivity loss parameter is recorded at several seismic hazard return periods. Several iterations are performed within each level. This provides the information for construction of interdependent fragility curves. Each interdependent fragility curve presents the effect that increasing the strength of coupling has on failure likelihood. This interdependent effect on the response occurs as a sudden phenomenon. Once the interdependency exists, the response is highly affected. However, increasingly larger values of the strength of coupling, though influential to the response, do not affect the network performance as dramatically as they do from independence to small interdependence.

The explanation for this sudden transition (which is also applicable to the sudden transition observed in the steep response of each infrastructure) is related to the fragmentation patterns of the networks. As the number of removed elements increases, the network fragments into several disconnected clusters. Among those clusters, there is one whose order is larger than the average order of the remaining clusters. That large cluster becomes the backbone to hold the network as a single entity. However, there is a critical moment in which a few nodes become essential to maintain the structure of the

system—or cluster. If those nodes are absent, the network rapidly partitions into several small islands.

The insights gained about interdependent failure mechanisms, response regimens, and role of nodes for topology and flow traversal, provide a strong basis for devising effective mitigation actions that are discussed in the following chapter.

CHAPTER 7

ENHANCEMENT OF NETWORK FUNCTIONALITY

Accounting for the interdependent effects in the response of networked systems accelerates degradation of their performance as the intensity of the disruption increases. If the networks are subjected to systematic removal of critical elements, their gradual deterioration has a marked peak on the interdependent effects—the absolute difference between the interdependent and independent response, normalized by the maximum observed independent response). Large extent simultaneous removals, such as the ones induced by earthquakes, proved that the interdependent effect can be as large as 100%. These significant effects on the response require devising mitigation strategies to control the propagation of disruptions within and across several infrastructures. However, interacting infrastructures are complex entities where the increasing intractability of disruption propagation obscures the identification of key network elements to enhance overall infrastructure response.

This chapter explores the effects that minimal interventions to enhance the performance of one infrastructure have on the response of another infrastructure. In particular, few elements of the power grid are improved either by lowering their centrality to flow traversal—decongesting them—or by reducing their seismic vulnerability. The effects of power grid enhancement on water network response are

monitored. Results reveal that key elements for power grid performance are not necessarily the key elements for maintaining multi-infrastructure flow transfer. The interface elements between infrastructures require their inclusion into the mitigation action plan. This implies that an effective mitigation action policy necessitates evaluation of the role that each network element of each infrastructure has in ensuring connectivity, facilitating flow traversal, and maintaining flow transfer at the interfaces.

This chapter introduces a straightforward interdependent rank-ordering procedure to account for the various responsibilities of network elements. It is conjectured that implementing mitigation actions on these elements will provide the necessary settings for propagating beneficial effects throughout the networks. This propagation is expected to occur the same way as cascading of detrimental effects spreads.

Finally, this chapter discusses a tool that can potentially be used to capture the required conditions for widespread propagation of disruptive effects among interdependent systems.

7.1 MINIMAL INTERVENTIONS

Evaluation of the effectiveness of competing mitigation actions on infrastructure systems requires exhaustive exploration of available alternatives. The same analysis approach used for estimating network performance after disruptions can be utilized to monitor network improvement after implementation of mitigation actions. Common methods for effective consequence minimization after disruptions rely upon optimization techniques. Researchers look for the optimal mitigation alternative to improve network service right after any disruption takes place. Optimal solutions are searched where they are

needed the most: in areas with high expected number of victims (i.e., fatalities and injuries). These areas are in general the same with low lifeline service reliability (Vanzi, 2000). Other studies focus on intensive search for the *critical* retrofit strategy. That strategy must result from trying an exhaustive set of retrofit alternatives, and selecting the one that produces the best possible improvement (Latora and Marchiori, 2005).

This study attempts to explore mitigation alternatives that simultaneously account for improvements in network flow patterns and in network topology stability. However, considering just topology, grid-like systems have a more redundant structure than sparse systems. Cycles within the network induce resilient response. Therefore, the retrofit strategy to first explore defines mitigation actions that locally modify network topology by adding cycles—in this case bypasses.

The objective is to induce a less vulnerable flow regime in network performance as measured by the connectivity loss, CL. The location for re-meshing the network is obtained from the vertex betweenness, B_v, rank-ordering. This sorting criterion assesses the amount of load or total flow that passes through the nodes at an unperturbed network state. The vertices to select for lessening their congestion are those where more than 65% of all possible flow paths pass through. Flow paths go from the generation set, F, to the distribution set, Q. For the power grid, being one of the core enabling systems for network interdependencies, its intact flow state indicates that elements P_2, P_5, and P_6 have the highest betweenness (Figure 7-1). Hence, these nodes are selected for remedial actions. The fundamental premise is to create bypasses around the most heavily loaded nodes to decongest them. Those new decongestion edges are constrained in their length due to edge expansion limitations. Growth of the United States high voltage transmission lines for the

2002 - 2011 period is projected to be 5% of the existing line-miles, despite the growth in demand of 20% (Bush, 2003).

Approximately 17% of the transmission lines of the power grid of this study are high voltage lines (e.g., > 350 kv). This leads to exploration of two network growth scenarios (i.e., bypasses), consistent with observed power grid growth trends. The first scenario corresponds to approximately 10 km (i.e., 5% growth) of new lines around element P_2 and partially around P_5; and the second scenario is approximately 40 km (i.e., 20% growth) of new lines around elements P_2, P_5, and P_6. Figure 7-2 presents an ideal bypassing mechanism to lessen the flow transfer burden on congested nodes.

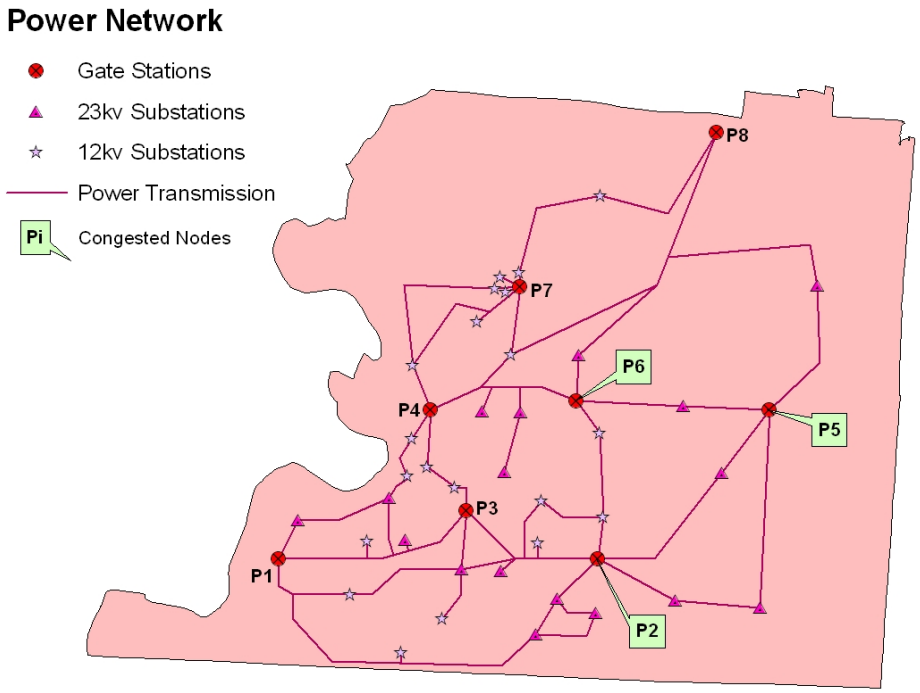


Figure 7-1. Congested nodes in the power grid as identified by their betweenness, B_v

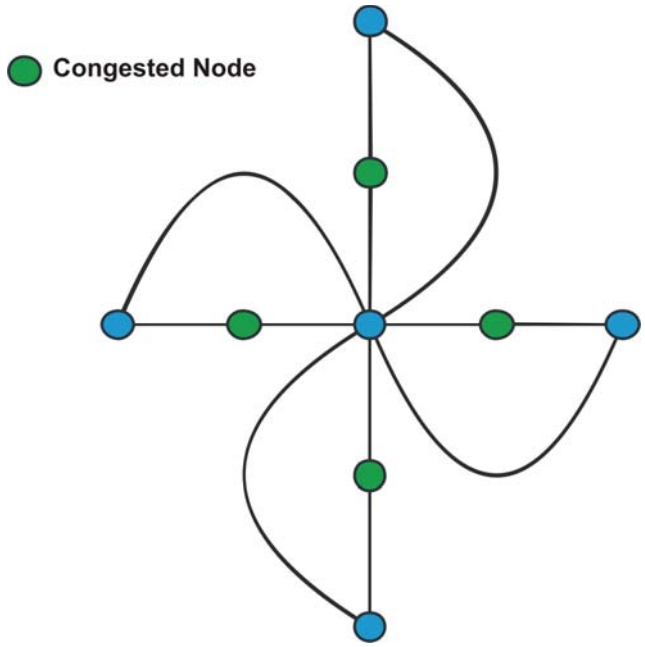


Figure 7-2. Ideal bypassing of congested nodes

Local mitigation actions increase the redundancy of networked systems. This is because new edges around congested vertices introduce additional cycles, which diversify the path portfolio used in flow transfer. To test this observation the network is subjected to a seismic hazard of 10% probability of exceedance in 50 years specified by ground motion contours. Figure 7-3 presents the impact of bypasses on the power grid for two transmission line growth levels. The response improves but not as significantly as expected. The steep fragility curves of the power grid have the explanation: for the seismic hazard under consideration (i.e., 475 year return period, which induces a PGA ~ 0.30g), the vulnerability of electric substations prevents new edges to be more effective, because failure occurs prematurely at most network nodes. This calls for seismic retrofit of vertices.

The effect of seismic retrofit is reported in the same figure. Improvements are introduced to the three identified vertices P_2, P_5, and P_6, which are high voltage electrical

substations. The fragility of an electrical substation can be reduced by providing adequate anchoring of its various subcomponents. It is assumed that the mean and the standard deviation of the LN (PGA) increase by 15% after anchoring of their subsystems. This results in a new median $PGA_m = 0.25$, and a new dispersion $\xi = 0.4$, as compared to the original $PGA_m = 0.20$, and $\xi = 0.35$ of Figure 6-30. Network improvements are more significant when seismic retrofit is implemented. However, this result provides an unfair view on the significance of node bypassing. For more modest disruptions such as lighting, vandalism, or aging, the nodes involved in the disruption can be effectively avoided through the new bypasses. They enhance local absorption of the perturbation. They contain it in a space that does not compromise large portions of the system.

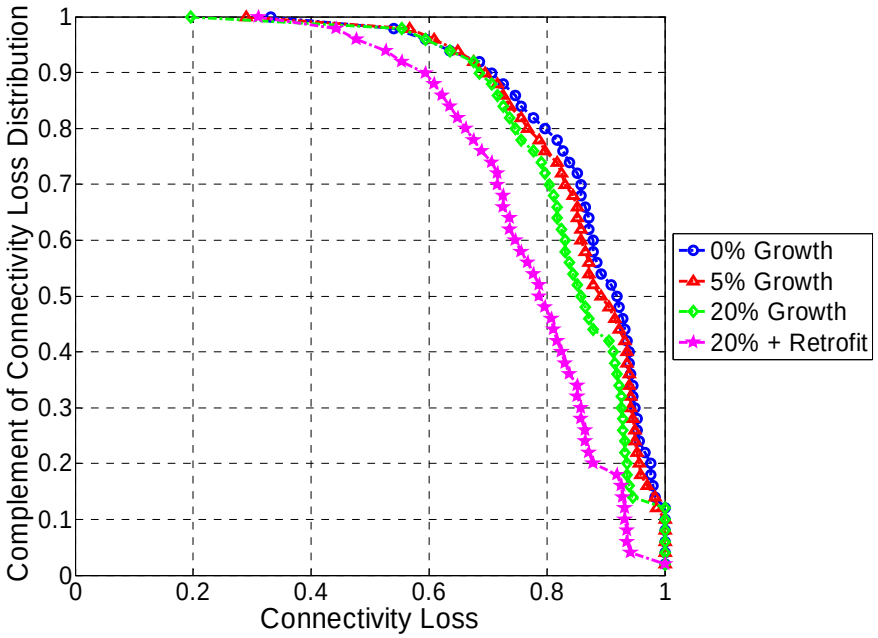


Figure 7-3. Impact of growth and seismic retrofit on power grid response

Power network growth for the seismic perturbation under consideration, is not sufficient to propagate visible benefits to the water network (Figure 7-4). The performance of the water system is reported when the interdependency is set to be $P(W_j|P_i) = 0.5$. If the power grid is retrofitted, then the effects on the water network response are more evident. This behavior of the water network is explained by the marginal contribution that construction of bypasses has when the networks are subjected to seismic hazards. However, for small rerouting needs, bypasses provide the solution. This indicates that a good mitigation policy must combine reduction of the vulnerability to external hazards, and alleviation of the internal congestion around funneling nodes.

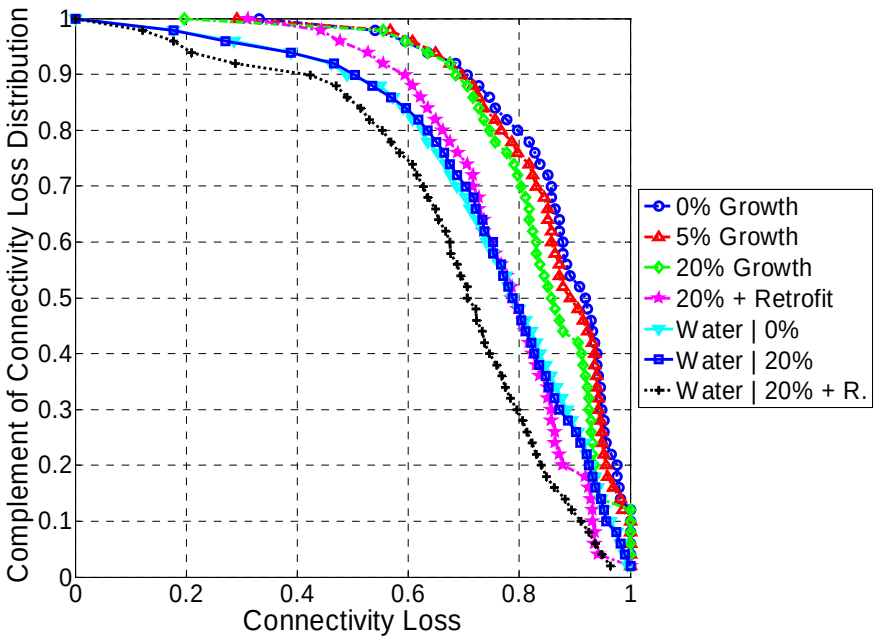


Figure 7-4. Impact of power grid growth and retrofit on water network response

An approach to better reduce the fragility of interdependent networks requires prioritization of the mitigation actions based upon a vertex rank-ordering that accounts for

the importance of each node to their multiple roles, and their susceptibility to multiple hazards. Each node has a place in network connectivity, flow traversal, flow transfer at the interface with other infrastructures, and network vulnerability. This concept can be generalized as a weighted average of vertex importance. If the possible roles that a node, v, can take for maintaining interdependent network performance are denoted by $H = 1, 2, \ldots, h$, then the formulation of the interdependent rank-ordering of a vertex, IRO_v, follows:

$$IRO_v = \frac{1}{h}\sum_{i=1}^{h} a_i v_i \qquad \textbf{(7-1)}$$

where a_i is the weighting factor given to each criterion of node importance. Sorting the IRO_v of all $v \in V(G)$ in ascending order provides a simple prioritization list to implement mitigation actions. Current research efforts are focused on finding optimal edge growth, independent network vertex retrofit, and interdependent interface vertex retrofit. This ultimate strategy is expected to propagate through the networks as an inverse avalanche in the same way some disruptions can generate cascading failures after exceeding a particular threshold of stability (Callaway et al., 2000; Watts, 2003).

7.2 CASCADING SUSCEPTIBILITY

The description of network response to external disruptions and to mitigation actions has been done in two different ways. One way systematically removes vertices according to a rank-ordered set which is updated after every removal. The other way simultaneously removes network nodes whose probability of failure, given certain hazard intensity, is

exceeded. The analysis of network response using these methods has provided valuable insights about their structure and function. However, looking ahead, an additional trend for analysis of complex systems may continue gaining acceptability. This trend relies upon definition of governing rules for each element of the interacting systems. These rules monitor the state of the nodes and update their state according to the conditions of their environment. If there is a flow overload, a node can absorb the event, or fail and let its neighbors to handle the peak. Sometimes the neighbors contain the event, but sometimes they cannot and the process propagates as an avalanche (Watts, 2002).

In the context of interdependent infrastructures, the overall state of their elements can be monitored by using few influential parameters. For instance, two network generic descriptors are the average susceptibility of failure (i.e., vulnerability), and the average importance for network functionality (i.e., centrality). As the intensity of the disruptions increases, or as the strength of the coupling among infrastructures increases, the networks show a critical value below and above which their state is completely different. This sudden transition is expected if the interdependent network problem is formulated as a set of interacting agents whose decisions are determined by the actions of their neighbors according to simple threshold rules.

Figure 7-5 depicts what would be a tool to determine the cascading susceptibility of interacting networks. Its axes represent average descriptors of the state the network. At any given time, if a disruption is induced, the interacting rules will allow the network to settle into another state. Each new state of the network will only depend on its current state—resembling a Markovian process. Exhaustive exploration of possible network states will indicate the conditions that lead to unstoppable large scale avalanches.

As an illustration, if the network is at a current state *P*, any perturbation able to induce an increase in average centrality or vulnerability will leave the network in a state that is closer to a precondition for cascading. Insights gained from this analysis will permit devising mitigation actions that inhibit the conditions that favor cascading development. Consequence minimization will be more effective since most of the disruptions would be managed at local scales.

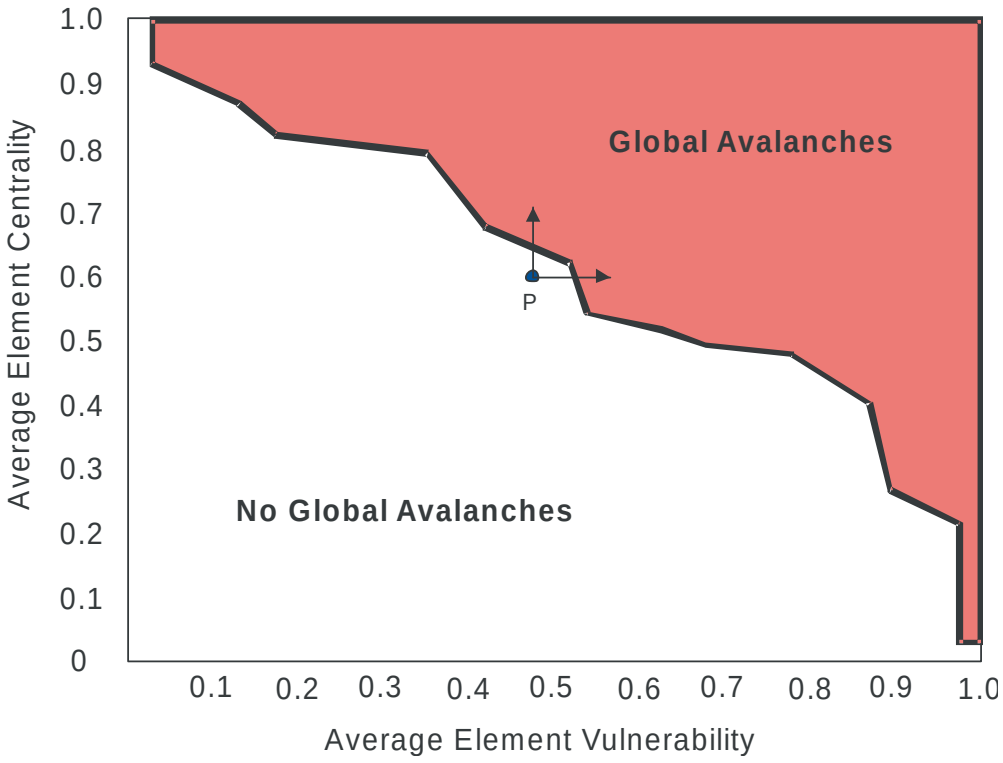


Figure 7-5. Conditions for interdependent network cascading failures

7.3 SUMMARY

The decision about which network elements should be the object of mitigation actions requires looking at the multi-infrastructure system as a single complex entity. For

the problem of seismic hazards, if mitigation actions are devised considering just the importance of the nodes in terms of their role to facilitate flow (i.e., lessen congestion), their impact is marginal. If seismic retrofit actions are implemented, then as expected, the effects of mitigation are more significant. However, if the external hazard were a less disruptive and common event, such as lightning, vandalism, or simple malfunctioning, decongesting loaded nodes with bypasses will allow the networks to effectively absorb the perturbation at a local scale.

Since the response of a network depends on the integrity of nodes and edges, keeping the most valuable elements will result in a network that, despite disturbances, is still able to provide its service. The importance of the nodes depends on their role for maintaining the topology of the system, for facilitating flow traversal, and for being instrumental in flow transfer at the interface with other networks. The contribution of the nodes to overall infrastructure vulnerability due to their fragility to external disruptions is another importance factor. This means that selecting nodes for mitigation actions in interdependent systems requires accounting for their multiple responsibilities.

This study introduces a simple prioritization scheme. The interdependent rank-ordering, IRO_v, of a vertex is defined as the weighted average of the ranking of each node in the various importance criteria. The nodes with the smallest *IRO* will be preferred for initiation of mitigation efforts. The claim here is that a cascade of beneficial effects can be propagated within the networks, in the same way that detrimental effects propagate as an avalanche.

More sophisticated mitigation actions require development of agent-based simulations. In these simulations, each node is an entity with governing rules whose

decisions are determined by the actions of their neighbors. Any disruption is either absorbed locally, or it grows and propagates throughout a significant portion of the networks. Inhibiting the preconditions for cascading failure will represent the most desirable mitigation policy.

CHAPTER 8

CONCLUSIONS AND FUTURE RESEARCH

The understanding of the structure and function of interdependent networks requires further exploration of the fundamental mechanisms that shape and condition their evolution. The topology of most infrastructures has distinct features that set them apart in terms of their ability to handle internal or external disruptions. The continuous functionality of critical infrastructures is taken for granted by the public. However, these networks are undergoing massive growth within their own systems, and across other systems. Interdependencies result in more efficient lifelines, but at the same time, the state of each network becomes more dependent on the state of other networks. This global interdependence creates a less tractable complex entity, which has the potential to induce unforeseen failure modes. The occurrence of devastating events efficiently propagates its detrimental effects through the network of physical infrastructures that sustain modern societies. Any infrastructure failure directly threatens the economy, the governance, the security, and the confidence of the population.

This chapter synthesizes the lessons learned from the exploration of the performance of interdependent systems. It indicates the significance of the work and its potential applications. Also, this chapter outlines some of the research tasks required to continue advancing the study of engineered networked systems.

8.1 CONCLUSIONS

Infrastructure systems are entities whose networked nature makes them suitable for rigorous mathematical analysis. Their description relies upon tools from graph theory, optimization, statistical physics, and structural engineering.

One of the first tasks to understand networks is to measure their properties according to their adjacency matrix—the array that describes which nodes are connected to others. This study describes properties that provide information about their global connectivity, local clustering, and overall shape. It also introduces a new parameter that allows quantification of the ability of the networks to reroute flow when there are localized perturbations. The parameter is referred to as the redundancy ratio, R_R. Networks that belong to the class of transmission and distribution of flows (e.g., electric energy, potable water, natural gas, and oil) exhibit redundancies close to the ones inherent in perfect square grids, where $R_R \sim 0.20$. Networks with abundant cycles—flow paths that start and end at the same node—exhibit more desirable properties, including a higher R_R. Complete graphs—networks where every node has a direct connection to every other node—have a redundancy ratio of $R_R = 1.0$.

In addition to the study of parameters that aid in the characterization and classification of networks according to their topological properties, there are parameters that help with the differentiation of the relative importance of network elements. Their rank-ordering ranges from purely topological criteria, to criteria that simultaneously account for topology and optimal flow patterns. This study introduces one criterion that captures optimal flow traversal within networks. It is referred to as the vertex

transshipment flow, W_v. It measures the ability of the nodes to facilitate flow transfer to meet supply and demand constraints. This parameter closely reproduces real flow phenomena, such as systems that are governed by physical laws of circuits in the case of power grids, or conservation of mass in the case of water networks.

Another set of metrics of interest are related with network performance. These parameters measure if the networks are able to provide their intended service. Additionally, these performance parameters can rely on various levels of input data, as in the previous case when the importance of the vertices of the networks was defined. Some only require topological information, while others necessitate more complete input data sets—including capacities, costs, supply, demands, etc. This study proposes a parameter that captures the ability of the network to meet the flow demands of its end-users. This parameter is referred to as the service flow reduction, S_{FR}, of the networks. This parameter can be directly related to the impact on the socio-economic apparatus in which the networks exist. This parameter can be measured in any infrastructure network, highlighting its usefulness for generic applications and for interdependent network analysis. Also, despite its more demanding input data set, it is still feasible for computer applications. In fact all of the parameters introduced in this research have polynomial-time algorithms—where the time to solve the problem scales as a polynomial function of the size of the problem, as opposed to other problems whose solution algorithms scale exponentially. This makes it desirable for integration with software development.

The availability of certain parameters to characterize network properties and response calls for the development of models that reproduce properties observed and measured in real systems. This study introduces a transmission and distribution network

model denoted as the TD model. The model is a realization from an aperiodic 2-lattice template with maximum vertex degree $\Delta(G) = 8$—a square grid with bypasses between all nodes. The properties of real power and water networks are well reproduced by the TD models. This provides a tool for abstract exploration of the performance of transmission and distribution networks—in the absence of complete data of real systems. It also provides the basis for the study of the impact of network size and order on their measured properties.

The explicit modeling of network interdependencies poses a challenge on the selection of the parameters that capture the phenomenon. There are several dimensions that induce coupling between infrastructure systems. They range from the environment in which the network operates, to the quality of the labor force that maintains and control their functionality. This study proposes a framework to more objectively capture the interdependence among infrastructures. Most of the interconnectedness is captured by a single entity referred to as the interdependent adjacency matrix, $A_{G|G'}$. This matrix accounts for the location of the interdependency, the direction of the interaction, and the representation of the strength of coupling, $P_{Gj|Gi'}$, which can be tuned from independence to interdependence. Geographical immediacy is the key aspect for defining the locations where the interfaces of the networks interact.

The response of networked systems is investigated for two types of disruptions: systematic removal of network elements—one at a time—and simultaneous removal of several elements. Systematic removals include deletion of elements according to their importance in terms of connectivity, flow traversal, and optimal flow distribution. Random selection is also included to represent spontaneous failure of elements. The

response of the networks to these disruptions indicates that the evolution of the detrimental effects—as a function of the amount of removed elements—is bounded by the random removal and the removal that targets the nodes with the largest vertex degree, $d(v)$. All response metrics, including topological measures that help in deciphering network failure modes, display the same response bounds. More interestingly, the response of real systems, as captured by the transshipment flow, Wv, is closer to the response exhibited by the random removal strategy. This implies that flow patterns within the networks make use of a highly diverse portfolio of paths to accomplish their intended function. The explanation of these observations is contained in the tails of the distribution function that describe the properties of the vertices. Differences between power and water network response are also attributable to their vertex property distribution—in particular to their vertex degree distribution.

Monitoring the response to systematic removal of interdependent networks also reveals the presence of performance bounds. The stronger the strength of coupling among networks the faster the degradation of their desirable properties. Also, the removal strategies are bounded by the random selection and the targeted vertex degrees. The quantification of the effect of interdependencies on network response reveals that their effects remain strong throughout most removal fractions. These effects are quantified using the interdependent effect, I_e, parameter, which measures the relative importance of the differences between interdependent and independent responses. The peaks of the I_e are always first observed at removal fractions less than 20%, precisely where the networks operate during normal and mildly abnormal conditions. Another interesting outcome of the interdependent effects is that when the removal of the

elements of the independent network is strategic (i.e., focusing on the important elements) then the effect on the dependent network is smaller than in the case where the removal of the independent network elements is at random. This contradicts the expected result: if one systematically disrupts a network in the most damaging possible way, one expects to see larger effects on other networks. The explanation of this contradiction is simple. The nodes that are critical to maintaining the integrity of one network are not necessarily the nodes that maintain the flow at the interface between networks.

The response of the networks to widespread initial failure, as induced by earthquakes, shows a particular type of behavior as the intensity of the hazard increases. Their observed probabilistic response resembles a step function. The response of these systems is monitored as their individual elements fail according to their fragility functions, which is an example of a random method. Consequences include an increase in the intensity of the hazard, and a simultaneous decrease in the amount of nodes and links in the network. There is a critical moment in which some of the remaining elements have the responsibility of connecting various fragments of the disrupted network. If they are absent, the entire network collapses. Such response resembles a phase transition which is a sudden change in state. This same sudden response is observed if the networks are either independent or interdependent. The effect of the strength of coupling also induces a sudden chance in state. The change in the response of the systems between independence and low interdependence is larger than the change between low strength of interdependence and complete interdependence. However, in both cases their relative impact on the independent response is substantial. This information is captured

in the interdependent fragility curves, which are introduced as a tool to enhance current loss estimation and risk assessment methods.

The regime at which the networks respond to seismic hazards changes as the intensity increases. As an example, the water network at low peak ground acceleration (PGA) exhibits a polynomial decay (i.e., linear in log-log scale) in the complement of the response parameter's cumulative density function (CDF). At a medium PGA, the complement of the CDF exhibits a linear decay, and at a high PGA, the complement distribution exhibits a logarithmic decay. The power grid shows logarithmic decays for all PGA values. This is due to the high fragility of its critical elements such as high voltage substations.

The insights gained about response distributions, the importance of elements, and the interdependent effects can be utilized to devise effective mitigation actions. These minimal interventions have the potential to induce significant improvements in network response. This study explores the ability of mitigation actions to decrease critical nodes' congestion and seismic vulnerability. Both are useful depending on the nature of the disruption. Small disruptions are controlled with bypasses and large disruptions are contained if their elements are less fragile. The results indicate that an effective mitigation policy should account for the simultaneous importance of every node to maintain network connectivity, to facilitate network flow, to be instrumental in flow transfer between networks at their coupling interfaces, and to reduce overall vulnerability. A method for interdependent rank-ordering is proposed. This study claims that by identifying the necessary interdependent elements, future corrective measures will ensure

propagation of benefits during their life cycle and throughout the same networked medium in which cascading failures propagate.

Finally, the search for mitigation actions also calls for more sophisticated representation of the transient dynamics of networked systems. Agent-based simulations may provide the tools to identify the conditions that lead to cascading failures. Having identified those conditions, one can take preventive actions which will ultimately ensure no widespread failure within or across infrastructures.

8.2 APPLICATIONS AND FUTURE RESEARCH

Immediate applications of the outcomes of this research are related with the enhancement of loss estimation methods. The algorithmic structure of the network characterization and response parameters is suitable for efficient software implementation. These methods will aid city planners, public officials, and utility owners in determining the best mitigation policies so that the response of their critical infrastructures is improved at the interdependent level. Results from an interdependent analysis provide clues about the topological deficiencies of the networks, about the characteristics of their interfaces with other networks, and about the expected magnitude of the interdependent effects. This type of analysis can also shed some light on common network failure modes, and on how to better plan network expansions.

Other applications, outside the earthquake engineering field, include implementation of the interdependent formulation in the design of expert systems to control the performance of military ships. These systems are finite entities where several infrastructures coexist, interact, and share the same spatial domain.

The understanding of complex systems is at its infancy. From the ongoing development of the present research several issues have arisen as topics for more in-depth study. A selected list of research activities include:

- development of models to capture the topology of telecommunication networks;
- investigation of the *scaling* of network response to disruptions—effect of network size on its functionality;
- exploration of network growth patterns to induce resilient architectures;
- development of interdependent systems using artificial network architectures (e.g., TD models, and telecommunication models);
- development of models to explore cascading susceptibility within interdependent infrastructures;
- investigation of reliability metrics to describe the likelihood of failure of interdependent systems subjected to multiple hazards and facilitate development of decision support tools for consequence minimization;
- development of models that capture the transient dynamics of the constituent networks;
- study of the impact of network state of operation on interdependent formulation; and
- formulation of the interdependent adjacency matrix as a multi-graph problem.

APPENDIX A: COMPUTATIONAL COMPLEXITY

An *algorithm* is a step-by-step procedure for solving a problem. An *instance* is a special case of a problem with data specified for all the problem parameters. The different steps an algorithm typically performs are (1) assignment steps (e.g., assigning a value to a variable), (2) arithmetic steps (e.g., addition, subtraction, multiplication and division), and (3) logical steps (e.g., comparison of two numbers).

A widely used approach to measure the performance of an algorithm is referred to as the *worst-case analysis* (Ahuja et al., 1993). This analysis, independent of the computing environment, provides an upper bound on the number of steps—and time—that a given algorithm can take on *any* problem instance. To express the time requirement of an algorithm, it is necessary to define a measure of the *complexity* of the problem instances. It is possible to express such a measure as a function of the problem size—which can be directly related to the network parameters n (i.e., order), m (i.e., size), C (i.e., largest cost), and U (i.e., largest capacity). A *time complexity function* for an algorithm is a function of the problem size and specifies the largest amount of time needed by the algorithm to solve any problem instance of a given size. This function measures the rate of growth in solution time as the problem size increases.

The "big O" notation has become common in computational mathematics, and replaces the lengthy expression "the algorithm required cnm time for some constant c" by the equivalent expression "the algorithm requires $O(nm)$ time." This means that the big

O notation indicates only the most dominant term in the running time, because for sufficiently large *n*, terms with a smaller growth rate become insignificant as compared to terms with a higher growth rate. In essence, this complexity measure states the asymptotic growth rate of the running time. Currently the scientific community considers a network algorithm "good" if its worst-case complexity is bounded by a polynomial function of the problem's parameters (i.e., a polynomial function f *n*, *m*, log *C*, and log *U*). Any such algorithm is said to be a *polynomial-time algorithm*. An algorithm is said to be an *exponential-time algorithm* if its worst-case running time grows as a function that cannot be polynomially bounded by the input length. Some examples of exponential time bounds are $O(nC)$, $O(2^n)$, $O(n!)$, and $O(n^{\log n})$.

A simple example to illustrate this complexity measure is the problem of changing the values of the non-zero elements of any asymmetric adjacency matrix. This problem will take time $O(m)$, because the assignment has to be done *m* times, where *m* represents the number of non-zero entries—or edges. Regarding network characterization parameters, the algorithm to calculate the mean distance *L* of a network takes time $O(nm)$ for an undirected network of *n* vertices and *m* edges (Ahuja et al., 1993). This is accomplished by running *n* times a *breadth-first search* algorithm. This algorithm starts at a particular node *v*, and from it step-by-step visits its neighbors, then the neighbors of its neighbors, and so on, until every edge and vertex is visited once, and scores are kept for the shortest distances. This operation takes time $O(n+m) = O(m)$. Doing this *n* times—for *n* vertices—the total worst-case running time is $O(nm)$.

The algorithm to calculate the clustering coefficient, γ, of a graph *G* runs in time $O(nm)$. The algorithm to compute the average degree, $d(G)$, of a graph *G*, is linear with

the number of vertices. Therefore it runs in time $O(n)$. The algorithm for calculating the redundancy ratio, R_R, of a graph G has a worst-case complexity of $O(n^3m)$. This comes from finding the number of node independent paths from a vertex v to all other vertices—which is true in a complete graph where $| V(\Gamma^2(v)) | = n$. This is accomplished using a conventional maximum flow - minimum cut algorithm that runs in time $O(n^2m)$ and is based upon successive shortest path algorithm solutions (Ahuja et al., 1993). Then repeating the process n times, one for each vertex $v \in V(G)$ as initial vertex, gives the bound of $O(n^3m)$. Approximate methods can establish the number of node independent paths between two vertices in linear time with respect to network size (White and Newman, 2001). Hence, an approximate solution to the global R_R can be run in time $O(n^2m)$. Table A-1 summarizes the computing time of the topological and performance parameters to characterize networked systems.

Table A-1. Worst case complexity of algorithms for network analysis

Network Property	Running Time
Mean distance, L	$O(nm)$
Modified mean distance, L'	$O(nm)$
Clustering coefficient, γ	$O(nm)$
Vertex degree, $d(v)$	$O(n)$
Redundancy ratio, R_R	$O(n^3m)$
Vertex degree rank-ordering	$O(n^2)$
Vertex betweenness rank-ordering	$O(n^4)$
Vertex transshipment rank-ordering	$O(n^2m^2 \log U)$
Vertex random selection	$O(n)$
Efficiency, E	$O(n^2m)$
Connectivity loss, C_L	$O(n^2m)$
Service flow reduction, S_{FR}	$O(n^2m^2 \log U)$

where,

m	graph size
n	graph order
U	largest upper bound of flow in $e \in E(G)$

APPENDIX B: NETWORK ALGORITHMS

B.1 MODIFIED MEAN DISTANCE, *L'*

FunctionB.1 [L'_1, L'_2, *DegSepU*] = ModifiedMeanDistance (*A*)

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Function:    Calculates reciprocal harmonic mean, L'
Details:     Serves as a global measure of network connectivity
Input:       Adjacency matrix of the network, A
Output:      Mean and median reciprocal harmonic mean L'1, and L'2, and matrix of shortest
             path lengths, DegSepU
Filename:    ModifiedMeanDistance.m
Date:        April 22, 2005
Author:      Leonardo Dueñas-Osorio
Revision:    0
Test:        TModifiedMeanDistance.m
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%
% 1 Initialization
%
DegSepU = PathLengthUndirected (A); % FunctionB.1.1
[a1,b1] = size(DegSepU);
MeanL = zeros(a1,1);
MedianL = zeros(a1,1);

%
% 2 Calculation of mean of means and median of medians
%
VectorOnes = ones(a1,1);
VectorZeroes = zeros(a1,1);
DegSepUInv = DegSepU;
DegSepUInv(find(eye(a1))) = VectorOnes;
DegSepUInv = 1./DegSepUInv;
DegSepUInv(find(eye(a1))) = VectorZeroes;
for a3 = 1 : a1
   [a4,b4,c4] = find(DegSepUInv(a3,:));
   [a5,b5] = size(c4);
   if b5 == 0
      MeanL(a3,1) = 0;
      MedianL(a3,1) = 0;
```

```
   else
      TempDegSepUInv = DegSepUInv(a3,:);
      TempDegSepUInv(:,a3) = [];
      MeanL(a3,1) = mean(TempDegSepUInv);
      MedianL(a3,1) = median(TempDegSepUInv);
      clear TempDegSepUInv;
   end
end
L1 = mean(MeanL);
L2 = median(MedianL);
```

FunctionB.1.1 *DegSepU* = PathLengthUndirected (*A*)

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```
Function: Computes shortest path lengths for undirected graphs with non-negative costs
Details: Uses Floyd-Warshal algorithm to generate shortest paths
Input: Valid adjacency matrix, *A*: (i) No self-loops, (ii) No Repetitions (iii) Symmetric
Output: Matrix indicating degree of separation among vertices, *DegSepU*
Filename: PathLengthUndirected.m
Date: September 20, 2004
Author: Leonardo Duenas-Osorio
Revision: 0
Test: TPathLengthUndirected.m
```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%
% 1 Variable Declaration
%
[n,m] = size(A);
DegSepU = zeros(n,n,2);
t=1;

%
% 2 Initialization
%
DegSepU(:,:,t) = A;
for i = 1 : n
   for j = i : m
      if ( i ~= j & DegSepU(i,j,t) == 0 )
         DegSepU(i,j,t) = inf;
         DegSepU(j,i,t) = DegSepU(i,j,t);
      end
   end
end

%
% 3 Evaluation of shortest paths
%
for t = 2 : n+1
   for i = 1 : n
```

```
      for j = i : n
         if ( j ~= t-1 )
            DegSepU(i,j,2) = min( DegSepU(i,j,1), DegSepU(i,t-1,1) + DegSepU(t-1,j,1) );
            DegSepU(j,i,2) = DegSepU(i,j,2);
         else
            DegSepU(i,j,2) = DegSepU(i,j,1);
            DegSepU(j,i,2) = DegSepU(i,j,1);
         end
      end
   end
   DegSepU(:,:,1) = DegSepU(:,:,2);
end
DegSepU(:,:,1) = [];
```

B.2 CLUSTERING COEFFICIENT, γ

FunctionB.2 *Gamma*1 = ClusteringCoefficient(*A*)

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```
Function: Calculates average clustering coefficient of networks
Details: Evaluates the number of connections among the neighbors of each vertex and normalizes it by the total possible number of connections among them
Input: Adjacency matrix, *A*
Output: Average clustering coefficient, *Gamma*1
Filename: ClusteringCoefficient.m
Date: May 28, 2004
Author: Leonardo Dueñas-Osorio
Revision: 0
Test: TClusteringCoefficient.m
```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

[n,n] = size(A);
Ne = zeros(n,n);
for i=1:n
   ii = 0;
   for j=1:n
      if A(i,j) > 0
         ii = ii+1;
         Ne(i,ii) = j;
      end
   end
end
for i=1:n
   for j=1:n
      Clust = zeros(n,n);
      for g=1:n
         if Ne(i,j) == 0 | Ne(i,g) == 0
```

```
        elseif A(Ne(i,j),Ne(i,g)) > 0
           Clust(j,g) = 1;
        end
      end
      Clu(j,1) = sum(Clust(j,:));  % Sums edges within neighbors of v that start at the j^th neighbor
   end
   K1 = length(find(A(i,:)));  % Implies symmetry
   Clus = sum(Clu)/2;
   if K1 == 0 | K1 == 1
      gamma1(i,1) = 0;
   else
      gamma1(i,1) = Clus / (K1*(K1-1)/2);
   end
end
Gamma1 = mean(gamma1);
```

B.3 VERTEX DEGREE, *d*(*G*)

FunctionB.3 [*De*, *pDes*, *K*1, *K*2] = VertexDegreeDistribution(*A*)

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

Function: Computes the vertex degree distribution of a network
Details: Rank-orders frequency of occuerrence
Input: Adjacency matrix, *A*
Output: Cumulative distribution of vertex degree, *pDes*, vertex degree vector, *De*, mean vertex degree, *K*1, and median vertex degree, *K*2
Filename: VertexDegreeDistribution.m
Date: April 22, 2005
Author: Leonardo Dueñas-Osorio
Revision: 1; Removed "for" loop
Test: TVertexDegreeDistribution.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

[n,n] = size(A);
Flag = zeros(n,n);
for i=1:n
   [r1,c1] = size(find(A(i,:)));
   De(i,1) = c1;
end
Des = sort(De);
K1 = (mean(Des));
K2 = (median(Des));
N = size(Des);
for i=1:N(1,1)
   p(i,1) = 1 - i/(N(1,1)+1);
end
pDes = [p, Des];
```

B.4 REDUNDANCY RATIO, R_R

FunctionB.4 [*RedundancyVector*, *RR*] = RedundancyMatrix(*A*)

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
Function:    Computes the average redundancy ratio of networks
Details:     Focuses on alternative paths connecting vertices with their second order
             neighboors
Input:       Valid adjacency matrix, A: (i) no self-loops, (ii) no Repetitions (iii) Symmetric
Output:      Vector indicating relative degree of redundancy, and average Redundancy, RR
Filename:    RedundancyMatrix.m
Date:        October 2, 2005
Author:      Leonardo Duenas-Osorio
Revision:    0
Test:        TRedundancyMatix.m
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%
% 1 Initialization
%
[n,n] = size(A);
Redundancy = zeros(n,n);
RedundancyVector = zeros(n,1);


%
% 2 Calculation
%
for i = 1 : n
   a1 = find(A(i,:));
   G2 = zeros(n,1);
   Neighborhood2 = zeros(n,1);
   for j = 1 : length(a1)
      a2 = find(A(a1(j),:));
      for k = 1 : length(a2)
         if i < a2(k) & G2(a2(k)) ~= 1
            Counter1 = 0;
            Connection = 0;
            MM = A;
            [Connection, Pathh] = ShortestPath1to1Sequence(MM, i, a2(k));
            % indicates if connection exist between i,j and provides the path sequence
            while Connection ~= Inf
               [r1,c1] = size(Pathh);
               if c1 == 2
                  MM(i,a2(k)) = 0;
                  MM(a2(k),i) = 0;
                  Counter1 = Counter1 + 1;
```

```
            else
               c2 = c1-1;
               Set = Pathh(2:c2);
               for m = 1 : (c1 - 2)
                  MM(Set(1,m),:) = zeros(1,n);
                  MM(:,Set(1,m)) = zeros(n,1);
               end
               Counter1 = Counter1 + 1;
            end
            [Connection, Pathh] = ShortestPath1to1Sequence(MM, i, a2(k));
         end
         Redundancy(i,a2(k)) = Counter1;
         Redundancy(a2(k),i) = Counter1;
         G2(a2(k)) = 1;
      end
   end
   Redundancy;
   N1 = a1;
   length(N1);
   N2 = find(Redundancy(i,:));
   Neighborhood2(N1) = 1;
   Neighborhood2(N2) = 1;
   if length(N1) == 0
      SizeG2 = 1;
   else
      SizeG2 = length(find(Neighborhood2 == 1));
   end
   RedundancyVector(i,1) = sum(Redundancy(i,:))/(SizeG2)^2;
end
RR = mean(RedundancyVector);
```

FunctionB.4.1 [*DegSep*, *Pathh*] = ShortestPath1to1Sequence(*A*, *i*, *j*)

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

Function: Computes shortest path between specified vertices and recovers path sequence
Details: Uses Dijkstra's one to all nonnegative costs algorithm for undirected graphs
Input: Valid adjacency matrix, *A*, initial node, *i*, and final node *j*
Output: Shortest path between a particular node and all other nodes, *DegSep*, path sequence between nodes *i* and *j*, *Pathh*
Filename: ShortestPath1to1Sequence.m
Date: October 2, 2005
Author: Leonardo Duenas-Osorio
Revision: 0
Test: TShortestPath1to1Sequence.m

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

%
% 1 Initialization
%
```

```
MM = A;
[N,N] = size(MM);
v = zeros(N,1);
d = zeros(N,1);
v(:,1) = Inf(N,1);
v(i,1) = 0;
Aux = ones(N,1);
Aux(i) = 0;
permanent = i;
CountDown = 0;

%
% 2 Calculation of shortest path
%
for q = 1 : N
   [r3,c3] = size(find(Aux));  % determines column of temporal nodes
   [r4,c4] = find(Aux);
   [r1,c1] = size(find(MM(permanent,:)));
   [r2,c2] = find(MM(permanent,:));
   vTemp = 0;
   if r3 == 0 | c1 == 0
     continue
   end
   Temporals = intersect(r4', c2)';
   [r7,c7] = size(Temporals);
   if r7 == 0
      Temporals = 0;
   end
   if r7 == 0  % updates permanent nodes even when no temporal nodes are available
      [r8,c8] = find(v(:,1)>=0);
      Admisibles = intersect(r4,r8);
      vTemp = v(Admisibles,1);
      NextPerm = min(vTemp);
      [r5,c5] = find(vTemp == NextPerm);
      Aux(r4(r5(1,1),1),1) = 0;
      permanent = r4(r5(1,1),1);
      continue
   else
      [r6,c6] = size(Temporals);
      for u = 1 : r6  % evaluates minimum distance and keeps track of path sequence
         d1 = v(Temporals(u,1),1);
         v(Temporals(u,1),1) = min( v(Temporals(u,1),1), v(permanent,1) + 1);
         d2 = v(Temporals(u,1),1);
         if d1 ~= d2
            d(Temporals(u,1),1) = permanent;
         end
      end
   end
   [r8,c8] = find(v(:,1));
   Admisibles = intersect(r4,r8);
   vTemp = v(Admisibles,1);
```

```
    NextPerm = min(vTemp);
    [r5,c5] = find(vTemp == NextPerm);
    Aux(r4(r5(1,1),1),1) = 0;
    permanent = r4(r5(1,1),1);
end

%
% 3 Recovery of path sequence
%
jj = j;
if jj == i
    Pathh = 0;
else
    Pathh(1,1) = j;
    for r = 2 : N
        if d(j) == 0
            Pathh = 0;
            break
        end
        Pathh(1,r) = d(j);
        j = d(j);
        if j == i
            break
        end
    end
end
DegSep = v(jj,1);
```

APPENDIX C: SCALING OF NETWORK SIZE

Table C-1. Scaling of m as a function of n

Network Type	Scaling of m
Complete graphs	$\frac{1}{2}\, n\, (n - 1)$
Square grids	$\frac{1}{2}\, d(v)\, [n - n^{0.5}]$
Open rings	$\frac{1}{2}\, d(v)\, (n - 1)$
Stars	$(n - 1)$
d-lattices	$\frac{1}{2}\, d(v)\, n$
Trees	$(n - 1)$
ER models	$\frac{1}{2}\, d(G)\, n$
WS models	$\frac{1}{2}\, d(v)\, n$
BA models	$(n - v_o)\, m_i + m_o$
TD substrates with $d(G) \sim 8$	$\frac{1}{2}\, d(v)\, [n - \frac{3}{2}\, n^{0.5}]$
TD model for power grids	$1.20\, n^{1.02}$
TD model for potable water networks	$1.32\, n^{1.02}$

where,

m	graph size
n	graph order
$d(v)$	degree of vertex v
$d(G)$	average vertex degree
v_o	initial number of nodes in BA growth
m_o	initial number of edges in BA growth
m_i	number of new links per new vertex

APPENDIX D: PHASE TRANSITION IN NETWORK RESPONSE

Starting with n nodes and adding edges at random until $m = n \times \ln(n)$ a phase transition in the number of connected nodes is observed. This effect is more notorious as n becomes larger. The number of edges, m, is twice the amount of edges required to guarantee that almost any random graph is connected (Erdös and Rényi, 1959). The data is normalized with respect to the order, n, and size, m, of each graph.

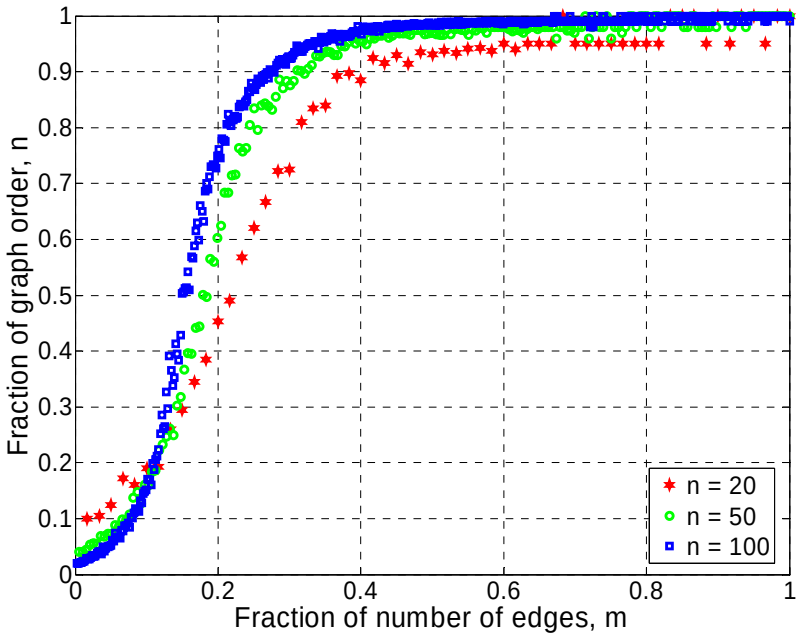


Figure D-1. Phase transition of the order of random graphs.

REFERENCES

Advanced Grounding Concepts (2005). WINIGS-F 2.3a: multi-phase power flow analysis of power systems. Atlanta, Georgia: Advanced Grounding Concepts.

Ahuja, Ravindra K., Thomas L. Magnanti and James B. Orlin (1993). Network flows: theory, algorithms, and applications. Upper Saddle River, New Jersey: Prentice Hall.

Albert, Réka, Hawoong Jeong and Albert-László Barabási (2000). "Error and attack tolerance of complex networks." Nature **406**: 378-382.

Albert, Réka and Albert-László Barabási (2002). "Statistical mechanics of complex networks." Reviews of Modern Physics **74**(1): 47-97.

Albert, Réka, István Albert and Gary L. Nakarado (2004). "Structural vulnerability of the North American power grid." Physical Review E **69**(2): 025103.

American Lifelines Alliance (2002). Seismic design and retrofit of piping systems. Washington D. C.: American Lifelines Alliance (ALA).

American Society of Civil Engineers. (2005). "ASCE responds to hurricane Katrina." ASCE disaster response programs. September 2005. Retrieved 6 September 2005 from <http://www.asce.org/static/hurricane>.

Amin, Massaoud (2000). "Toward self-healing infrastructure systems." IEEE Computer **33**(8): 44-53.

Amin, Massaoud (2001). "Toward self-healing energy infrastructure systems." IEEE Computer Applications in Power **14**(1): 20-28.

Applied Technology Council (1991). ATC-25 Seismic vulnerability and impact of disruption on lifelines in the coterminous United States. Redwood City, California: Applied Technology Council (ATC).

Axelrod, Robert and Michael D. Cohen (1999). Harnessing complexity: organizational implication of scientific frontier. New York, NY: The Free Press.

Banavar, Jayanth R., Amos Maritan and Andrea Rinaldo (1999). "Size and form in efficient transportation networks." Nature **399**: 130-132.

Barabási, Albert-László (2003). Linked: how everything is connected to everything else and what it means for business, science, and everyday life. New York, NY: Plume.

Barabási, Albert-László and Eric Bonabeau (2003). "Scale-free networks." Scientific American **288**(5): 50-59.

Baran, Paul. (1964). "Introduction to distributed communication networks." Memorandum RM-3420-PR. January 2004. The RAND Corporation. Retrieved 20 August 2005 from <http://www.rand.org/publications/RM/baran.list.html>.

Bollobás, Béla (1998a). Random graphs. Modern graph theory. Ed. New York, NY: Springer**:** 215-252.

Bollobás, Béla (1998b). Modern graph theory. New York, NY: Springer.

Borgatti, Stephen P. (2005). "Centrality and network flow." Social Networks **27**: 55-71.

Bush, R. (2003). "The future of transmission in North America." Transmission and Distribution World **5**(1).

Cagnan, Z. and R. Davidson (2003). Post-earthquake lifeline service restoration modeling. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 255-264.

Callaway, Duncan S., Mark E. J. Newman, Steven H. Strogatz and Duncan J. Watts (2000). "Network robustness and fragility." Physical Review Letters **85**(25): 5468.

Chang, Stephanie E., Hope A. Seligson and Ronald T. Eguchi (1996). Estimation of the economic impact of multiple lifeline disruption: Memphis light, gas, and water division case study. Technical Report No. NCEER-96-0011. Buffalo, New York: Multidisciplinary Center for Earthquake Engineering Research (MCEER).

Chang, Stephanie E. and Nobuoto Nojima (2001). "Measuring post-disaster transportation system performance: the 1995 Kobe earthquake in comparative perspective." Transportation Research Part A **35**: 475-494.

Chang, Stephanie E. and Hope A. Seligson (2003). Evaluating mitigation of urban infrastructure systems: application to the Los Angeles department of water and power. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 474-483.

Comfort, Louis K., Milos Hauskresht and Jeen Sheng Lin (2004). Dynamic networks: modeling change in environments exposed to risk. Annual Research Conference of the Association of Public Policy and Management. Atlanta, GA.

Crucitti, Paolo, Vito Latora, Massimo Marchiori and Andrea Rapisarda (2003). "Efficiency of scale-free networks: error and attack tolerance." Physica A **320**: 622-642.

Crucitti, Paolo, Vito Latora and Massimo Marchiori (2004). "Model for cascading failures in complex networks." Physical Review E **69**(4): 045104.

Dastous, Jean Bernard and André Filiatrault (2003). Seismic displacement at interconnection points of substation equipment. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 597-606.

Diestel, Reinhard (2000). Graph Theory. 2nd ed. New York, NY: Springer.

Dueñas-Osorio, Leonardo, James I. Craig and Barry G. Goodno (2004). Probabilistic response of interdependent infrastructure networks. 2nd annual meting of the Asian-pacific network of centers for earthquake engineering research (ANCER). Honolulu, Hawaii. July 28-30.

Dueñas-Osorio, Leonardo, James I. Craig and Barry G. Goodno (2005a). "Optimal flow approach to quantify performance of networked systems." Manuscript to be sent to Physical Review E.

Dueñas-Osorio, Leonardo, James I. Craig and Barry G. Goodno (2005b). "Seismic response of critical interdependent networks." Manuscript under review by the IAEE Journal of Earthquake Engineering and Structural Dynamics.

Dueñas-Osorio, Leonardo, James I. Craig and Barry G. Goodno (2005c). "Interdependent response of networked systems." Manuscript under review by the ASCE Journal of Infrastructure Systems.

Eguchi, Ronald T., James D. Goltz, Craig E. Taylor, Stephanie E. Chang, Paul J. Florez, Laurie A. Johnson, Hope A. Seligston and Nail C. Blais (1998). "Direct economic losses in the Northridge earthquake: a three-year post-event perspective." Earthquake Spectra **14**(2): 245-264.

Eguchi, Ronald, T. and Douglas G. Honegger (2003). Standard guidelines to assess the seismic fragility of water transmission systems. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 504-511.

Erdös, Paul and Alfred Rényi (1959). "On random graphs." Publicationes Methematicae **6**: 290-297.

Federal Emergency Management Agency (2005). HAZUS-MH MR1 Technical Manual. Washington D. C.: Federal Emergency Management Agency (FEMA).

Frankel, A., S. M. Petersen, C. Mueller, K. Haller, R. Wheeler, E. V. Leyendecker, R. Wesson, S. C. Harmnsen, C. H. Cramer, D. M. Perkins and K. S. Ruckstales (2002). Documentation for the 2002 update of the national seismic hazard maps, Open-file report 02-420. United States Geological Survey.

French, Steven P. and Xudong Jia (1997). "Estimating societal impacts of infrastructure damage with GIS." URISA Journal **9**(1): 31-43.

Friedman, Thomas L. (2005). The world is flat: a brief history of the twenty-first century. 1st edition. New York, NY: Farrar, Straus and Giroux.

Gell-Man, Murray (1994). Complex adaptive systems. Complexity: metaphors, models, and reality. Ed. G. Cowan, D. Pines and D. Meltzer. Westview Press.

Goh, K. I., B. Kahng and D. Kim (2001). "Universal behavior of load distribution in scale-free networks." Physical Review Letters **87**(27): 278701.

Graf, W. P., C. E. Taylor, J. H. Wiggins, L. Lund and T. Voltz (2003). Guidelines for defining natural hazards performance objectives for water systems. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 455-464.

Guimera, Roger, Alex Arenas, Albert Díaz-Guilera and Francesc Giralt (2002). "Dynamical properties of model communication networks." Physical Review E **66**(2): 026704.

Haimes, Yacov Y. and Pu Jiang (2001). "Leontief-based model of risk in complex interconnected infrastructures." Journal of Infrastructure Systems **7**(1): 1-12.

Haimes, Yacov Y., Barry M. Horowitz, James H. Lambert, Joost R. Santos, Kenneth G. Crowther and Chenyang Lian (2005a). "Inoperability input-output model for interdependent infrastructure sectors. II: case studies." Journal of Infrastructure Systems **11**(2): 80-92.

Haimes, Yacov Y., Barry M. Horowitz, James H. lambert, Joost R. Santos, Chenyang Lian and Kenneth G. Crowther (2005b). "Inoperability input-output model for interdependent infrastructure sectors. I: theory and methodology." Journal of Infrastructure Systems **11**(2): 67-79.

Hatestad (2004). Computer applications in hydraulic engineering: connecting theory to practice. Sixth ed. Hatestad Press.

Hauer, John F. and Jeff E. Dagle (1999). White paper on review of recent reliability issues and system events. Pacific Northwest National Laboratory.

Heller, Miriam (2001). "Interdependencies in civil infrastructure systems." The Bridge **31**(4): 6p.

Holme, Peter and Beon Jun Kim (2002). "Growing scale-free networks with tunable clustering." Physical Review E **65**(2): 026107.

Holme, Peter, Beon Jun Kim, Chang No Yoon and Seung Kee Han (2002). "Attack vulnerability of complex networks." Physical Review E **65**(5): 056109.

Huyck, Charles K., Ronald T. Eguchi, Reid M. Watkins, Hope A. Seligson, Stephen Bucknam and Edward Bortugno (2003). URAMP (Utilities Regional Assessment of Mitigation Properties): a benefit-cost analysis tool for water, wastewater and drainage utilities. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 484-493.

Hwang, Howard H. M., Huijie Lin and Masanobu Shinozuka (1998). "Seismic performance assessment of water delivery systems." Journal of Infrastructure Systems **4**(3): 118-125.

Institute of Electrical and Electronics Engineers (1997). Recommended practices for seismic design of substations, IEEE-693 Standard. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers (IEEE) Standards Department.

Institute of Electrical and Electronics Engineers (2004). Guide for electric power distribution reliability indices, IEEE-1366 Standard. New York, New York: Institute of Electrical and Electronics Engineers (IEEE) Standards Department.

Jain, Sanjay and Sandeep Krishna (2002). "Crashes, recoveries, and "core shifts" in a model of evolving networks." Physical Review E **65**(2): 026103.

Kaiser, Marcus and Claus C. Hilgetag (2004). "Spatial growth of real networks." Physical Review E **69**(3): 036103.

Kondrasuk, Jack N. (2004). "The effects of 9-11 and terrorism on human resource management: recovery, reconsideration and renewal." Employee Responsibilities and Rights Journal **16**(1): 11.

Kosterev, Dimitry N., Carson W. Taylor and William A. Mittelstad (1999). "Model validation for the August 10th, 1996 WSCC system outage." IEEE Transactions on Power Systems **14**(3): 967-979.

Latora, Vito and Massimo Marchiori (2001). "Efficient behavior of small-world networks." Physical Review Letters **87**(19): 198701.

Latora, Vito and Massimo Marchiori (2002). "Is the Boston subway a small-world network?" Physica A **314**: 109-113.

Latora, Vito and Massimo Marchiori (2005). "Vulnerability and protection of infrastructure networks." Physical Review E **71**(1): 015103.

Li, Hong-Nan, Wen-Long Shi and Su-Yang Wang (2003). Simplified seismic calculation method for coupled system of transmission lines and their supporting tower. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 697-706.

Li, Jie and Jun He (2002). "A recursive decomposition algorithm for network seismic reliability evaluation." Earthquake Engineering and Structural Dynamics **31**: 1525-1539.

Little, Richard G. (2002). "Controlling cascading failures: understanding the vulnerabilities of interconnected infrastructures." Journal of Urban Technology **9**(1): 109-123.

Lleras-Echeverri, German and Mauricio Sánchez-Silva (2001). "Vulnerability analysis of highway networks, methodology and case study." Trasport **174**(4): 223-230.

Matt, Howard and André Filiatrault (2003). Seismic response of high voltage transformers. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 647-656.

Mid-America Earthquake Center (2004). MAEViz introduction and tutorial. Urbana, Illinois: Mid-America Earthquake (MAE) Center.

Milgram, Stanley (1967). "The small world problem." Psychology Today **2**: 60-67.

Moreno, Y., R. Pastor-Satorras, A. Vázquez and A. Vespignani (2003). "Critical load and congestion instabilities in scale-free networks." Europhysics Letters **62**(2): 292-298.

Motter, Adilson E. and Ying-Cheng Lai (2002). "Cascade-based attacks on complex networks." Physical Review E **66**(6): 065102.

Motter, Adilson E., Takashi Nishikawa and Ying-Cheng Lai (2002). "Range-based attack on links in scale-free networks: are long-range links responsible for the small-world phenomenon?" Physical Review E **66**(6): 065103.

Multidisciplinary Center for Earthquake Engineering Research (1998). Engineering and Socioeconomic impacts of earthquakes: an analysis of electricity lifeline disruptions in the New Madrid area. M. Shinozuka, A. Rose and R. T. Eguchi. Buffalo, NY, Multidisciplinary Center for Earthquake Engineering Research (MCEER)**:** 189p.

Nakamura, Ikuo (2003). "Characterization of topological structure on complex networks." Physical Review E **68**(4): 045104.

National Research Council (2002). Making the nation safer: the role of science and technology in countering terrorism. Washington D. C.: The National Academies Press.

Newman, D. E., Bertrand Nkei, B. A. Carreras, Ian Dobson, V. E. Lynch and Paul Gradney (2005). Risk assessment in complex interacting infrastructure systems. Thirty eight Hawaii International Conference on System Science. Hawaii. January.

Newman, Mark E. J. and Duncan J. Watts (1999). "Scaling and percolation in the small-world network model." Physical Review E **60**(6): 7332.

Newman, Mark E. J. (2001a). "Scientific collaboration networks. I. Network construction and fundamental results." Physical Review E **64**(1): 016131.

Newman, Mark E. J. (2001b). "Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality." Physical Review E **64**(1): 016132.

Newman, Mark E. J., Steven H. Strogatz and Duncan J. Watts (2001). "Random graphs with arbitrary degree distributions and their applications." Physical Review E **64**(2): 026118.

Newman, Mark E. J. (2003). "The structure and function of complex networks." Society for Industrial and Applied Mechanics Review **45**(2): 167-256.

Newman, Mark E. J. (2005). "A measure of betweenness centrality based on random walks." Social Networks **27**: 39-45.

Nojima, Nobuoto and Massata Sugito (2003). Development of a probabilistic assessment model for post-earthquake residual capacity of utility lifeline systems. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 707-716.

O'Rourke, Michael and Gustavo Ayala (1993). "Pipeline damage due to wave propagation." Journal of Geotechnical Engineering **119**(9): 1490-1498.

Ostrom, Dennis K. (2003). System Earthquake Risk Assessment (SERA II). Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 587-596.

Park, Juyong and Mark E. J. Newman (2004). "Statistical mechanics of networks." Physical Review E **70**(6): 066117.

Posner, Richard, A. (2005). "Our incompetent government." The New Republic **233**(20): 23-27.

Quimpo, Rafael G. and Sue-Jen Wu (1997). "Condition assessment of water supply infrastructure." Journal of Infrastructure Systems **3**(1): 15-22.

Rardin, Ronald L. (2000). Optimization in operations research. Upper Saddle River, NJ: Prentice Hall.

Riley, Michael J., Leon Kempner Jr. and Wendelin H. Mueller III (2003). A comparison of seismic (dynamic) and static load cases for electric transmission structures. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 687-696.

Rinaldi, Steven, M., James P. Peerenboom and Terrence K. Kelly (2001). "Identifying, understanding, and analyzing critical infrastructure interdependencies." IEEE Control Systems Magazine **21**(6): 11-25.

Sachtjen, M. L., B. A. Carreras and V. E. Lynch (2000). "Disturbances in a power transmission system." Physical Review E **61**(5): 4877.

Schiff, Anshel J. and Leon Kempner Jr. (2003). Issues and guidance for IEEE 693 equipment qualification tests. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 607-616.

Sen, Parongama, Subinay Dasgupta, Arnab Chatterjee, P. A. Sreeram, G. Mukherjee and S. S. Manna (2003). "Small-world properties of the Indian railway network." Physical Review E **67**(3): 036106.

Shargel, Benjamin, Hiroki Sayama, Irving R. Epstein and Yaneer Bar-Yam (2003). "Optimization of robustness and connectivity in complex networks." Physical Review Letters **90**(6): 068701.

Shumuta, Yoshiharu (2003). Seismic risk management system for electric power facilities. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 677-686.

Song, Junho, Armen Der Kiureghian and Jerome L. Sackman (2003). Interaction between electrical substation equipment connected by rigid bus slider. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 617-626.

Strogatz, Steven H. (2001). "Exploring complex networks." Nature **410**: 268-276.

The Economist (2005). "The shaming of America." Economist **376**(8443): 11.

The White House (2003). The national strategy for the physical protection of critical infrastructures and key assets. Washington D. C.: Government Printing Office.

United States - Canada power system outage task force (2004). Final report on the August 14th, 2003 blackout in the United States and Canada: causes and recommendations. United States Department of Energy and Canadian Department of Natural Resources.

United States Census Bureau. (2002). "Census 1990." November 2004. United States Census Bureau. Retrieved 2 October 2005 from <http://www.census.gov/main/www/cen1990.html>.

United States Department of Energy (2000). The changing structure of the electric power industry: an update. Washington D. C.: Energy Information Administration.

Vanzi, Ivo (2000). "Structural upgrading strategy for electric power networks under seismic action." Earthquake Engineering and Structural Dynamics **29**: 1053-1073.

Walker, David M. (2002). "9/11: the implications for public sector-sector management." Public Administration Review **62**(s1): 94-97.

Watts, Duncan J. and Steven H. Strogatz (1998). "Collective dynamics of small-world networks." Nature **393**(4): 440-442.

Watts, Duncan J. (1999). Small worlds: the dynamics of networks between order and randomness. Princeton, New Jersey: Princeton University Press.

Watts, Duncan J. (2002). "A simple model of global cascades on random networks." Proceedings of the National Academy of Sciences **99**(9): 5766-5771.

Watts, Duncan J. (2003). Six degrees: the science of a connected age. New York, NY: W. W. Norton & Company.

Werner, Stuart D., Craig E. Taylor and James E. Moore II (1997). "Loss estimation due to seismic risks to highway systems." Earthquake Spectra **13**(4): 585-604.

White, Douglas R. and Mark E. J. Newman (2001). "Fast approximation algorithms for finding node-independent paths in networks." Santa Fe Institute Working Papers **01-07-035**.

Yang, Han and Sun Shaoping (2003). Seismic reliability of urban network systems. Sixth U.S. conference and workshop on Lifeline earthquake engineering, August 10-13. Ed. J. E. Beavers. Long Beach, California: American Society of Civil Engineers. 445-454.

Zhang, Peengcheng, Srinivas Peeta and Terry Friesz (2005). "Dynamic game theoretic model of multi-layer infrastructure networks." Networks and Spatial Economics **5**(2): 147-178.

Zhao, Liang, Kwangho Park and Ying-Cheng Lai (2004). "Attack vulnerability of scale-free networks due to cascading breakdown." Physical Review E **70**(3): 035101.

Zhao, Liang, Ying-Cheng Lai, Kwangho Park and Nong Ye (2005). "Onset of traffic congestion in complex networks." Physical Review E **71**(2): 026125.

Zhou, Shi and Raúl J. Mondragón (2004). "Accurately modeling the internet topology." Physical Review E **70**(6): 066108.

Zimmerman, Ray D., Carlos E. Murillo-Sánchez and Deqiang Gan. (2005). "MatPower: A Matlab power system simulation package." February 2005. Power Systems Engineering Research Center. Retrieved 20 September 2005 from <http://www.pserc.cornell.edu/matpower>.

VITA

Leonardo Augusto Dueñas Osorio was born in Bogotá, Colombia in 1976. He graduated as a Civil Engineer from the Universidad de La Salle in Bogotá in 1996. He pursued his graduate studies at the master level in structural and earthquake engineering at the Universidad de Los Andes in Bogotá. Starting in 1998, Leonardo led projects related to seismic design of railroad infrastructure, seismic design of mid-rise concrete structures, and quality control of engineering practice with the Italian consulting firm, Società Tecnica Internazionale. During the same period, he completed studies in project management at the Universidad Javeriana in Bogotá. During the fall of 2000, Leonardo enrolled as a graduate student at the Massachusetts Institute of Technology, where he would earn a master's degree with research focus on high performance structures. In the fall of 2001, he joined the group of doctoral students of the School of Civil and Environmental Engineering at the Georgia Institute of Technology.