Defining Research and Development Directions for Modeling and Simulation of Complex, Interdependent Adaptive Infrastructures

Robert J Glass, Walter E Beyeler, Stephen H Conrad, Nancy S Brodsky, Paul G Kaplan, and Theresa J Brown

Infrastructure Complexity R&D Group

National Infrastructure Simulation and Analysis Center
Sandia National Laboratories
Albuquerque, New Mexico 87185

Acknowledgement

The National Infrastructure Simulation and Analysis Center supported this work. NISAC's core partners are Sandia National Laboratories and Los Alamos National Laboratory. Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000. Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405 ENG-36.

Executive Summary

National and economic security and indeed, the quality of life in the U.S., depend on the continuous, reliable operation of a complex set of infrastructures. The National Infrastructure Simulation and Analysis Center, or NISAC, provides advanced modeling and simulation capabilities for the analysis of critical infrastructures, their interdependencies, vulnerabilities, and complexities. These capabilities help improve the robustness of our nation's critical infrastructures by aiding decision makers in the areas of policy analysis, investment and mitigation planning, education and training, and near real-time assistance to crisis response organizations. The work of the Infrastructure Complexity Research and Development (R&D) Group is one of NISAC's long-term investments in understanding infrastructures and their interdependencies. Our charter and mandate call for identifying theories, methods, and analytical tools from the study of complex systems so that they can be brought to bear on critical infrastructure problems.

In this past year, the Infrastructure Complexity R&D Group has focused on filtering concepts from the Complex Systems literature that cross disciplines, and identifying study/modeling approaches that are of direct use to NISAC. During the course of this study, a description of infrastructures naturally emerged that is similar to the way complex systems have been described in physics, biology, economics, and other diverse areas where Complexity Science has provided fruitful insights. We also worked to identify common unresolved issues pertaining to the application of Complex Adaptive Systems modeling to real-world situations. We then defined our R&D effort for Complex, Interdependent Adaptive Infrastructure to resolve the issues of greatest importance to NISAC. For our R&D effort we address: *Process*, or how we can conduct this research to most effectively develop and apply analytical insights; *Technical Objectives*, or which technical questions most urgently need answering; and *Modeling Tools*, or which formulations show the most promise for providing answers. We briefly describe each of these below.

Process: We will continue study to identify promising concepts, approaches and modeling methods. Disseminating that information to NISAC technical staff, model developers, and model users will be accomplished through regular open discussions.

Technical Objectives: Technical objectives will be defined in phases such that the results of earlier phases are used to define the specific technical objectives of later phases. For our first phase, we have defined a set of specific technical objectives. These objectives are to conduct parametric studies with appropriate Complex Systems models that will:

- Fill in gaps in the current understanding of how simple transition rules and network geometries can combine to permit system-spanning cascades;
- Discover the ranges of parameter values that cause transitions in macro-scale behavior (e.g. between isolated failure and pervasive failure);
- Identify the influence of departures from common assumptions in current models such as network type and the critical influence of heterogeneity.

Modeling Tools: Because of their flexibility and the ready availability of agent modeling tools, we anticipate using agent based models to formulate and answer many of our specific research questions.

Table of contents

Executive summary

1. Introduction 4
   1.1 What is Complexity Science and How Does It Relate to Infrastructures? 4
   1.2 Under the NISAC Umbrella – Investigations of Complex Systems 6
   1.3 Roadmap to the Rest of this Document 8
2. Synthesis of Complexity Science 8
   2.1 Summary of Concepts 9
      2.1.1 Complex Networks 10
      2.1.2 Complex Behavior 13
      2.1.3 Growth, Evolution and Adaptation: Complex Adaptive Systems 16
   2.2 Summary of Model Technologies: 18
   2.3 Infrastructures as Complex Adaptive Systems: 18
3. Basic Research Issues 20
   3.1 Model and Network Abstraction 20
   3.2 Model and Network Behavioral Definition and Categorization 21
   3.3 Hazard/Problem and Solution/Fix Categorization 21
   3.4 Adaption/Evolution/Directed-Evolution 22
   3.5 Predictability 22
4. Definition of the R&D effort 23
   4.1 Process 23
   4.2 Technical Objectives 24
   4.3 Modeling Tools 25
5. References 26

Appendixes

Appendix A: Our Course of Study
Appendix B: Poster Presentation at NISAC's Capability Demonstration
Appendix C: An Illustrative Example of Complex Behavior on Complex Networks: The "Generic Sand Pile"

List of Figures

Figure 1. NISAC technical components and information flows 7
Figure 2. Illustration of natural and constructed network systems. From Strogatz [2001]. 11
Figure 3. Examples of idealized networks. (From Left to Right – Top row: ordered networks with differing degrees; Bottom row: random network, blended ordered-random network, and scale free network.) Note that the top right is fully connected and so it will have a path length of 1. From Strogatz [2001] 12
Figure 4. Example time series and frequency distribution of cascade size in the sandpile model. From Bak, Tang and Wiesenfeld [1988] 14
Figure 5. Elements and interactions in a complex adaptive network. 17

1. Introduction

National and economic security and indeed, the quality of life in the U.S., depend upon the continuous, reliable operation of a complex set infrastructures that includes electric power, oil and natural gas, transportation, water, communications, banking and finance, emergency services, law enforcement, government continuity, agriculture, and health services. Each infrastructure is very complicated, formed from a large number of subcomponents, connected in myriad ways. Each incorporates people who make decisions at scales from the individual, to cliques, to companies, to consortiums and larger groups. These infrastructures are made interdependent by complex and often poorly-understood linkages. These interdependencies allow disruptions in any single infrastructure to jeopardize the continuous operation of the entire system of infrastructures.

Understanding individual infrastructures and their complex interdependencies and vulnerabilities is essential for implementing effective policy for the enduring operation, regulation, and defense of the national infrastructure as a whole. This understanding requires the development of advanced modeling, simulation, and analysis capabilities. These capabilities are embodied within the National Infrastructure Simulation and Analysis Center (NISAC). A subset of these capabilities is created by the work of the Infrastructure Complexity Research and Development (R&D)Group. In this first section of our report "Defining Research and Development Directions for Complex, Interdependent Adaptive Infrastructures", we first introduce Complexity Science and how its study may help us better understand the functioning of systems of infrastructures (Section 1.1). We then put our work in context of NISAC, our charter, and how our network complexity investigations support them (Section 1.2). Finally, we end our introduction with a roadmap to the remainder of our report (Section 1.3).

1.1 What is Complexity Science and How Does It Relate to Infrastructures?

Complexity Science has been used to explore commonalities among events as varied as:

- Earthquakes
- Mass extinctions
- Major wars
- Traffic jams
- Major forest fires
- Epidemics
- Revolutions
- Landslides
- Stock market crashes
- Major power outages

All of these events have something in common – although we are unable to fully explain their causes nor predict their precise occurrences and magnitudes, they exhibit behaviors characteristic of systems that are "complex". In general, complex systems are composed of many interacting parts with simple rules of behavior. One finds that these systems often yield behavior that is not intuitively obvious at the outset, that *the whole is greater than the sum of the parts*. Because of this, complex systems are particularly resistant to investigation using the reductionist approach common to many scientific and engineering investigations in which detailed study of the system components is sufficient to understand the system as a whole.

In recent years, a general theory for complex systems has emerged that suggests there is a natural tendency for diverse complex systems to "self-organize" into what is called the "critical state", a

state of instability often described as being at the "edge of order and chaos". In such a state, cascading events of all sizes can occur at any time and thus are unpredictable except through measures of their statistics. The behavior (e.g., the propensity to cascade) and resiliency (e.g., attack vs. error tolerance) of the complex system has also been found to depend on the statistical characteristics of complex networks. Additionally, there is a growing realization that many such systems adapt, especially when people or biological processes are integral to the system, and thus are aptly described as "Complex Adaptive Systems" or CAS. Here, the two aspects of complex systems, their behavior and underlying network structure, are intertwined with feedbacks that cause the system to evolve. Research on CAS has found that networks evolving within one "network ecology" can be particularly susceptible to disruption when the nature of the threats changes.

In the context of Complexity Science, we may ask ourselves questions about how interdependent infrastructures respond to disruptions, such as:

- How can seemingly small initiating events (e.g., single point equipment failures) cascade into large infrastructure network disruptions?
- Is it always straightforward to identify the critical nodes in a system and protect them from failure?
- How does the structure of the connectivity between nodes affect network stability?
- Can we develop improved indicators of an infrastructure's status?
- How can we use simulations of networks abstracted from real infrastructures to look for unintended consequences of proposed policy?
- Are there general lessons to be learned about infrastructure networks that can be applied across many systems obviating study of each infrastructure in excruciating detail?

Let us consider an infrastructure as a network of nodes, connected to each other by links through which some form of material or information flows. Nodes could be:

- Power plants, transformers, power grid loads
- Computers and routers on the internet
- Institutions in a financial network
- Transportation hubs (airports)
- Telecommunications hubs
- People (individuals or groups) in a social network

The geometric configurations, or topologies, of these networks can be further abstracted to allow systematic study of the more general or generic infrastructure. We can define simple abstracted rules for node behavior as well as rules for the interaction of one node with another on the abstracted network. The abstract infrastructure is now entirely analogous to those studied in Complexity Science.

The provocative findings of Complexity Science concerning cascading failures on the one hand and topological resiliency on the other raise questions regarding possible *inherent* susceptibilities to collapse, or easily exploited weaknesses in infrastructures that arise from simple rules for node dynamics or from the infrastructure topology. Depending on the answers, a strategy of identifying and selectively protecting "critical nodes" may ultimately prove to be unavailing and a more nuanced approach for evolving robust infrastructures might be indicated. Additionally, infrastructures change over time; system behavior and system structure are inherently linked and evolve through adaptive feedback. Complexity Science suggests that scale-free networks that have evolved in a "network ecology" optimized for a tolerance to random node outages (characteristic of water, electricity, natural gas and other distribution systems) are particularly susceptible to directed attacks. As another example, consider the current "business ecology" where market liberalization encourages leanness and imposes pressures on key infrastructures to

cut overheads (often by building out redundancy). Based on principles uncovered in the context of Complexity Science, policies that encourage infrastructure efficiency during normal operations may make these infrastructures less robust in response to disruptions.

1.2 Under the NISAC Umbrella – Investigations of Complex Systems

NISAC provides advanced modeling and simulation capabilities for the analysis of critical infrastructures, their interdependencies, vulnerabilities, and complexities. These capabilities will help improve the robustness of our nation's critical infrastructures by aiding decision makers in the areas of policy analysis, investment and mitigation planning, education and training, and near real-time assistance to crisis response organizations.

The White House National Strategy for Homeland Security specifically cites the need for state-of-the-art, high-end modeling and simulation of the nation's critical infrastructures, and directs the Department of Homeland Security (DHS) to take NISAC as its foundation for these efforts. In fact, NISAC is one of the 22 organizations that compose the new department. NISAC was formally chartered in the USA Patriot Act of 2001 (October 26, 2001), in which Congress tasked NISAC "to serve as a source of national competence to address critical infrastructure protection and continuity through support for activities related to counter terrorism, threat assessment, and risk mitigation." More specifically, the Act states that NISAC support will include modeling, simulation, and analysis of critical infrastructure systems, to enhance understanding of their large-scale complexity and facilitate system modifications to mitigate threats and enhance the stability of critical infrastructures. In short, our mandate is to integrate modeling, simulation, and analysis into national infrastructure and asset protection planning and decision support activities.

Within this framework, the mandate for complexity investigations has been to conduct research on complex system theory and modeling so as to identify useful methodologies and tools. We are charged with developing and testing new models, simulation and analysis tools and evaluating their utility. We are to evaluate models and data related to infrastructure interdependencies and system evolution. This includes conducting research to anticipate evolutionary trajectories of infrastructures and their interactions. The ultimate goal of this work is to identify potential limitations and vulnerabilities arising from patterns of infrastructure evolution to allow development of more robust systems.

Figure 1 shows some of the ways in which the various technical components that make up Sandia's NISAC effort relate to one another. Our R&D for Complex, Interdependent Adaptive Infrastructures falls within the center box of the figure: *Generalized Infrastructure Network Investigations*. The term "generalized investigations" distinguishes this R&D from the less abstracted, more applied work in the surrounding boxes that focuses on specific systems and problems. However, there is symmetry and tight coupling between the generalized and specific components and indeed, the generalized investigations form a central and binding role.

Both our work and the Economic Consequence Modeling use an agent-based approach to examine the flow of goods and information through social and physical networks. Economic Consequence Modeling is focused on representing real networks and addressing specific applied problems that are posed. The generalized complexity investigations deal with more idealized networks and transition rules in attempting to elucidate generalized behaviors common to many infrastructure networks. The tight coupling results as we look to the agent based modeling to compare our idealized networks to real network topologies. Do we see similar behaviors in both

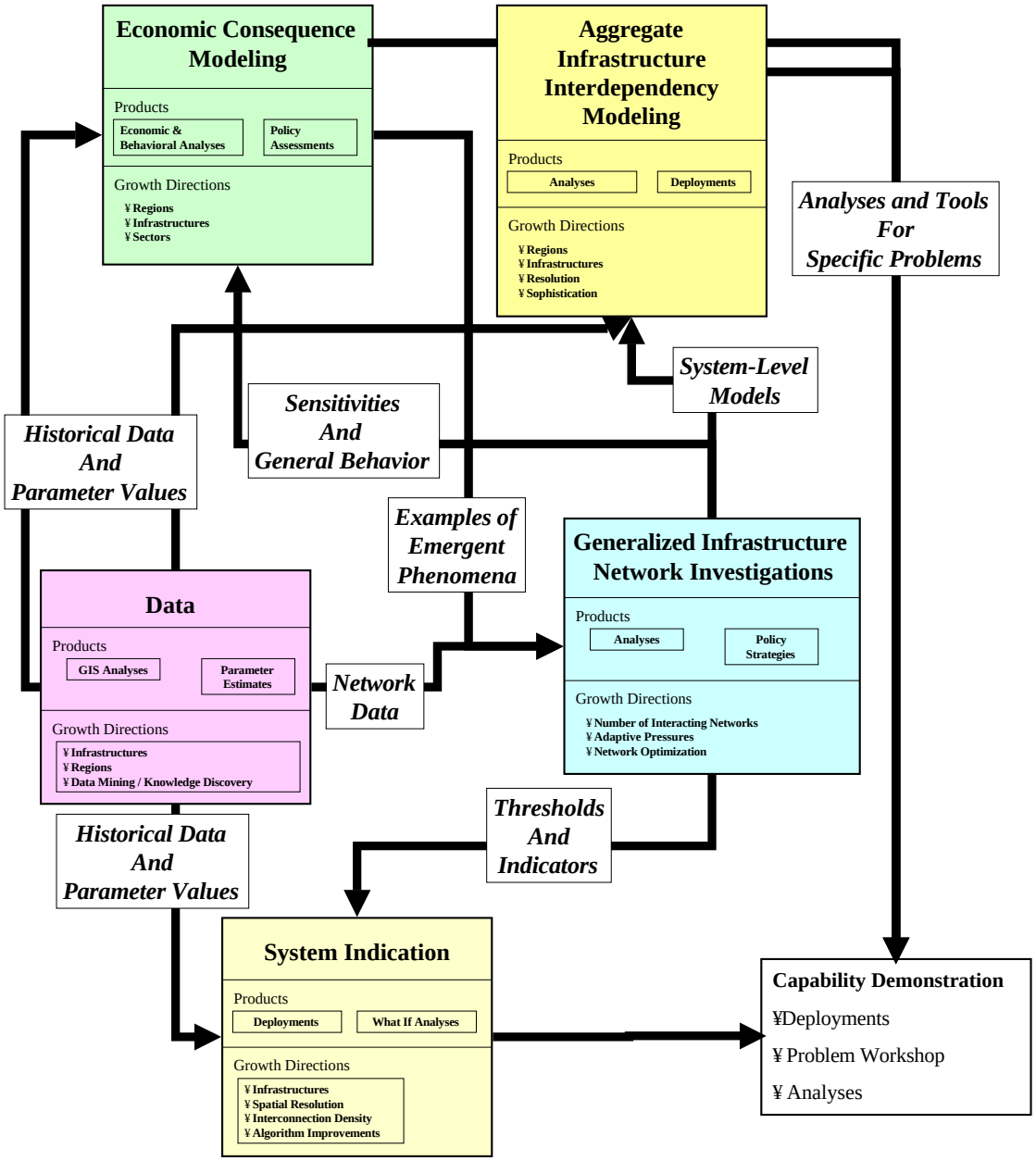


Figure 1. NISAC technical components and information flows

the abstracted and the more realistic networks? The Economic Consequence Modeling group will look to us to help better understand observed emergent behaviors, as we are able to perform systematic parametric analyses not easily performed in a much more highly representational model.

There are important ties to other NISAC technical components. We look to other sources to gather real network topologies for analysis and comparison. The highly general results from our work may inform the work of the other technical groups. For example, a better understanding of thresholds indicating a propensity for system instability will be useful in the System Indications component. This understanding could be translated into an improved operational algorithm for quantifying an infrastructure's status. Our work will uncover relationships between micro-conditions and macro-behavior, which can be used Aggregate Infrastructure Interdependency Modeling.

1.3 Roadmap to the Rest of this Document

The concepts introduced above will be discussed in greater detail in the remainder of this report. Section 2 summarizes the basic concepts from complexity research that we believe to give insight into the behavior of interdependent infrastructures. We discuss these concepts and their relevance for infrastructure modeling, and identify some of the limitations and gaps with respect to our goals. We find that interdependent infrastructures can be beneficially viewed as complex adaptive systems, and that this perspective can give us new insights into the effectiveness and limitations of actions we might take to sustain and protect infrastructures. Section 3 identifies and organizes some basic outstanding issues in more detail, and the areas in which further research should be directed. In Section 4, we describe a research agenda for next year that is designed to fill in the most significant gaps in our current understanding of infrastructures as Complex Interdependent Adaptive Systems. This report is further supported in a set of three Appendixes. Our course of study is outlined in Appendix A. Our Poster presentation at NISAC's capabilities demonstration in Portand and Seattle summarized some of our work, understanding, and future directions. We have included our poster in Appendix B. Also find a description of our demonstration modeling in Appendix C. A live, interactive version of this model was part of our Poster presentation.

2. Synthesis of Complexity Science

We reviewed the Complexity science literature, a diverse literature extending from the very qualitative to the highly theoretical. Because of the breadth and volume of the literature our review was necessarily selective. We concentrated on primary papers that are extensively cited, and survey articles that summarize the state of knowledge. We focused on filtering relevant concepts that crossed disciplines and identifying study/modeling approaches that would be useful to NISAC. See Appendix A where our course of study is described.

A description of infrastructures naturally emerged that is similar to the way complex systems have been described in physics, biology, economics, and other areas where this outlook has provided fruitful insights. Generally, these systems can be conceptualized as a set of 'nodes' that have certain rules of behavior and rules of interaction with other nodes. Interactions define a network of node connections, similar to an existing physical network (such as power transmission lines) or an informal social network along which information flows. Nodes form a population, which can be homogeneous or heterogeneous. Heterogeneous populations can be broken down into sub-classes that behave differently (i.e., have different rules of behavior and interaction) and have different but possibly overlapping interaction networks. The behavioral and interaction rules as well as the interaction network can change over time. This change can be random, directed, or shaped by feedbacks from system performance. In the last case the system can be thought of as adapting or evolving to improve performance.

In the following, we synthesize our review and illustrate the value of viewing infrastructures as complex systems. We first summarize concepts (Section 2.1), models (Section 2.2), and then focus on the implications for infrastructures and their modeling (Section 2.3).

2.1 Summary of Concepts

Research in complex systems has engendered many ideas from a variety of different points of view and across many scientific fields. These concepts include: percolation, chaos, self-similarity, fractals, self-organization, criticality, self-organized criticality (SOC), highly optimized tolerance (HOT), graph/network topology, social embeddedness ('herd' behavior), and adaptation. Complexity Science uses all of these concepts in a variety of contexts to both understand and model evolving chemistries, societies, intelligence, and life itself.

From an historical perspective, complexity is often considered an outgrowth of the field of nonlinear dynamical systems. Despite their apparent simplicity, models of nonlinear dynamical systems can exhibit behavior that is surprisingly irregular. This behavior discouraged study before the widespread availability of computers, which can produce numerical approximations of solutions. Lorenz's [1963] pioneering study of atmospheric flow identified surprising properties of certain seemingly simple models: they can produce outputs that never "settle" into a regular pattern, and they can amplify small changes in their initial conditions. Subsequently such systems, termed *chaotic*, have been extensively studied, and have been described in many popular technical books [e.g., Gleick, 1987]. A significant conclusion from this work is that even simple dynamical systems (which is a very broad class, including all of the systems that concern us) can, under some circumstances, be *unpredictable* in the sense that any error or "noise" in the description of the system properties, however small, will eventually cause the real system behavior to increasingly depart from the modeled behavior over time. This finding limits the kinds of conclusions that we can reasonably draw using models of dynamical systems.

In the mid 1980's it came to be recognized, however, that while deterministic Chaos explained many aspects of unpredictability for simple systems, it lacked the potency to explain unpredictability in systems that are highly complicated [e.g., Waldrop, 1992]. Deterministic chaos generates 'white noise', that is, it contains no temporal correlation. On the other hand, many highly complicated systems exhibit a time signature that is '1/f noise-like', i.e. that has a specific temporal correlation described by an inverse relationship between event size and frequency [e.g., Bak, 1996]. Such behavior is fractal. In addition, highly complicated systems have some behavioral similarities with systems at the critical point and thus concepts from the theories of Percolation [e.g., Stauffer and Anarony, 1992] and Critical Phenomena [e.g., Binney et al., 1992] have relevance. Thus, the concept of 'Complex Systems' rather than simply chaotic systems, and the field of 'Complexity Science' were born.

Here an integrative overview of complexity is provided. Subsequent sections are organized around the ideas with the greatest power to illuminate infrastructure behavior: Complex Networks (Section 2.1.1); Complex Behavior (Section 2.1.2); and Growth, Evolution and Adaptation: Complex Adaptive Systems (Section 2.1.3). These sections cover three essential aspects of infrastructure systems, respectively: their geometry; the macro or system scale dynamics within a fixed geometry; and the changes that their geometry and dynamics can undergo over time.

2.1.1 Complex Networks

Networks are a wonderfully flexible abstraction, and are generally applicable to the study of the wide variety of interacting systems that compose interdependent infrastructures, including the physical components, control systems, associated information exchanges, and the social and organizational systems that control and plan infrastructures. Consider the abstraction of an infrastructure into a series of nodes that are connected to each other by some form of interaction. The connections or links define the interactions that can (or do) occur between pairs of nodes. Connections might represent:

- Power transmission and distribution lines
- Information exchange
- Pipelines
- Communications lines
- Hyperlinks between internet resources
- Roadways
- Shared set of environmental conditions

The generality of abstract networks can be seen from the diverse examples of natural and constructed systems (see **Figure 2**) that have been described and analyzed using the concept. These examples include:

- Transportation systems
- Telecommunications
- Power grids
- Internet
- World Wide Web
- Social networks
- Ecological food webs
- Supply chains
- Metabolic pathways within cells
- Chemical reactions

While the behavior of real systems that we describe as networks depends on the particular processes that occur in nodes and along connections, many interesting properties have been found to arise just from the pattern of connections in the network. What's more, some important global properties of networks (such as whether they are subject to cascade failure) can be inferred from the statistical properties of network connections; an exact description of the network topology may not be required. For this reason, both real networks and abstract networks are often characterized by certain key statistics. Commonly specified characteristics of networks include:

- ***Degree distribution***: The *degree* of a node is simply the number of connections it has to other nodes. The *degree distribution* for a network describes the relative number of nodes with a given degree. Often the average degree is an important factor in determining global network properties.
- ***Clustering*:** The density of node interconnection can by characterized by a clustering coefficient. One common definition is the fraction of a node's neighbors (along connections) that are also connected to one another. The average value over the network is often cited.

- ***Path Length*:** The path length between two nodes is the number of connections along the shortest connecting path. A common summary measure for the network is the average path length over all pairs of nodes.

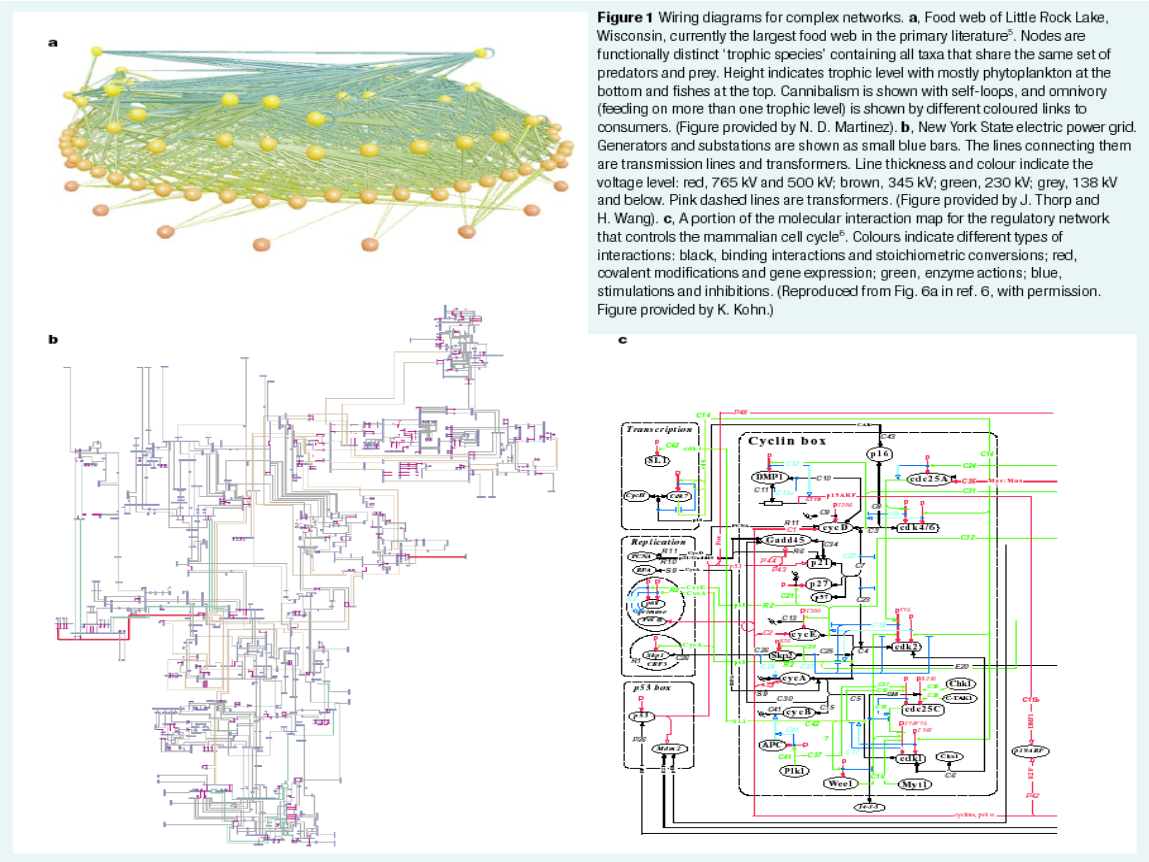


Figure 1 Wiring diagrams for complex networks. **a**, Food web of Little Rock Lake, Wisconsin, currently the largest food web in the primary literature[5]. Nodes are functionally distinct 'trophic species' containing all taxa that share the same set of predators and prey. Height indicates trophic level with mostly phytoplankton at the bottom and fishes at the top. Cannibalism is shown with self-loops, and omnivory (feeding on more than one trophic level) is shown by different coloured links to consumers. (Figure provided by N. D. Martinez). **b**, New York State electric power grid. Generators and substations are shown as small blue bars. The lines connecting them are transmission lines and transformers. Line thickness and colour indicate the voltage level: red, 765 kV and 500 kV; brown, 345 kV; green, 230 kV; grey, 138 kV and below. Pink dashed lines are transformers. (Figure provided by J. Thorp and H. Wang). **c**, A portion of the molecular interaction map for the regulatory network that controls the mammalian cell cycle[6]. Colours indicate different types of interactions: black, binding interactions and stoichiometric conversions; red, covalent modifications and gene expression; green, enzyme actions; blue, stimulations and inhibitions. (Reproduced from Fig. 6a in ref. 6, with permission. Figure provided by K. Kohn.)

Figure 2. Illustration of natural and constructed network systems. From Strogatz [2001].

Measurements of degree distribution, clustering, and path length for a wide variety of networks have been summarized by Albert and Barabasi [2002] and by Dorogovtsev and Mendes [2002]. However, that studies that have measured complex networks to date are few, the data underlying the measurements have varying degrees of accuracy, and the estimation error for some of the statistics is quite large. There is still no consensus with respect to the range in these measures/characteristics that complex networks may cover.

The topology of abstract networks is a result of the way they are constructed. For some networks, the construction process can include some stochastic elements. Network topologies studied in the theoretical literature often belong to a specific idealized class. Differences in generation processes lead to different network structures, and each class has a characteristic range of values for the statistics discussed above. Examples of the most common classes are shown in **Figure 3**.

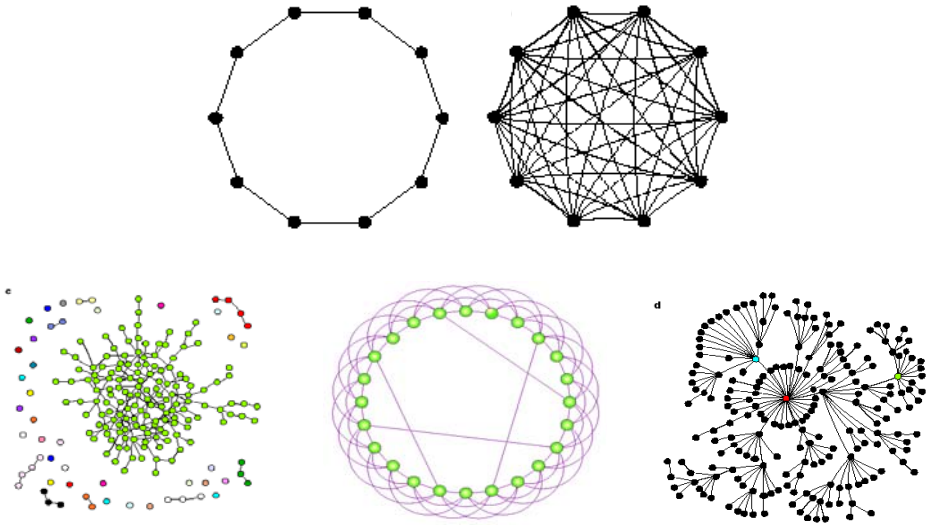


Figure 3. Examples of idealized networks. (From Left to Right – Top row: ordered networks with differing degrees; Bottom row: random network, blended ordered-random network, and scale free network.) Note that the top right is fully connected and so it will have a path length of 1. From Strogatz [2001].

- ***Ordered*** networks have a regular pattern of local connections, which is the same throughout the network. Examples include ring networks and regular grids or lattices. Such networks have a single degree, path lengths that are very long (on the order of the network size), and are highly clustered.
- ***Random*** networks are at the other end of the spectrum. Here a set of nodes is randomly interconnected: two nodes are selected at random and a connection is created between them. This process is repeated to achieve a certain number of edges. The study of such random networks began many years ago with the work of Erdos and Renyi [1959] who founded the branch of mathematics called Graph theory. Random networks have a degree distribution that is Poisson. They have very short path lengths, but are not highly clustered.
- ***Blended ordered-random*** networks combine features of ordered and random networks. Ordered and entirely random graphs/networks, while mathematically tractable, are often not very representative of the real networks in which we are interested. Blended ordered-random networks, which were first proposed by Watts and Strogatz [1998, WS model], can be generated from ordered networks by adding random links. Like many real networks, they tend to be highly clustered yet have short path lengths. Degree distributions span the range from single valued to Poisson.
- ***Scale free*** or ***fractal*** networks were introduced by Barabasi and Albert [1999, BA model] in response to observations of an incremental growth process that may reflect the way some real networks expand. When nodes are added to the network they preferentially connect to existing nodes that have high connectivity. This preferential connection seems to typify citations of technical papers, creation of hyperlinks in the World Wide Web, and other kinds of connections. The resulting networks are highly clustered, have short path lengths, and exhibit power law degree distributions.

Complex networks tend to have the following: a degree distribution that is often a power law, a high degree of clustering, and a small path length. This final attribute is often referred to as the

‘Small World’ characteristic after the work of Milgram [1967]. Networks with Small World characteristics, i.e., short path lengths, can span the spectrum from a single value of degree (ordered, fully connected) to Poisson (random), to somewhere between (blended ordered to random WS model), to power-law distributions (BA scale-free model) with some very influential or well-connected nodes. Thus small worlds can be all the way from ‘egalitarian’ to ‘aristocratic’.

Different classes of networks can show important differences in performance. For example, Albert, Jeong, and Barabasi [2000] compared the failure tolerances of random and scale-free networks. They found that random networks are more vulnerable to random node failure than scale-free networks. On the other hand, scale-free networks are more vulnerable to a directed attack that targets high-degree nodes.

2.1.2 Complex Behavior

Interesting conclusions about the performance and vulnerability of networks follow simply from their geometry. However, infrastructures are dynamic systems, i.e., they exhibit behavior. Across a wide variety of systems (e.g., the earth’s crust, ecosystems, markets, human conflict, the Internet) one finds strikingly similar behaviors: long-range spatial-temporal correlation, cascading events, and power law event size distributions. In other words, *they exude temporal-spatial fractals*. While by no means unanimously agreed upon within the literature, these behavioral signatures have grown to be diagnostic of *complex systems* and we will use them to define *complex behavior*.

From whence does *ubiquitous* complex behavior come? Consider a general system composed of a large number of interacting components such as electrons, molecules, grains, blocks, organisms, tribes, societies, or nations, each of which locally exchanges forces, energy, information, or in some way interacts. The system exists within an external environment (force field, environmental change, etc) that pushes or drives it. Macro-scale system behavior will result… what will it be? Is there some typical behavior shared by large classes of systems, or will the behavior always depend crucially on the details of each system? In the late 1980’s, Bak, Tang and Wiesenfeld [1987], whom we will refer to as BTW, made a rather extraordinary claim; they hypothesized that systems composed of many interacting components would exhibit a general behavior in that all such dynamical systems would *organize themselves* into a state with a complex but general structure characterized by *temporal-spatial fractal**s*** such as found at *critical points*. They named this behavior *Self-Organized Criticality* or *SOC*. After its introduction by BTW, SOC became a candidate for a general theory of Complexity [see Waldrop 1992]. SOC combines concepts of *Self-Organization* with *Critical Behavior* to explain *Complex Behavior* [Jensen, 1998].

- ***Self-organization***: The ability of certain nonequilibrium systems to develop structures and patterns in the absence of control or manipulation by an external agent (see Nicolis, 1989), e.g., patterns in chemical reactions (BZ reaction), fluid systems (convective cells), maybe even the development of structure in biological systems (slime molds).
- ***Criticality***: Has a precise meaning in equilibrium thermodynamics (see Binney et al., 1992). At a phase transition, the effect of a local perturbation dies off algebraically, whereas away from the transition point, it dies off exponentially. This means that even though nearest neighbors interact only locally, a local disturbance at the critical point will propagate throughout the entire system. It is in this sense that the system is ‘critical’, that all members of the system influence each other. In equilibrium systems, the system must be *tuned* to the critical point.

The essential idea of SOC is that the dynamical system drives itself (i.e., self-organizes) into a critical state (characterized by algebraic correlations), i.e., *no tuning is required*. For a system to evolve into an SOC dynamical state, a *separation of time scales* is required such that the processes that drive the system are slow compared to the internal relaxation processes. Earthquakes are a clear example: they occur over very short time scales, but are driven by stresses that accumulate over many years. The separation of time scales is intertwined with the existence of *thresholds* and *metastability*. Metastable states are ones in which the system is 'stuck' away from the stable, lowest energy state. When a system moves from a metastable state, it moves through a number of marginally stable states and then 'sticks' in another metastable state.

BTW's simple illustrative system was composed of identical elements connected on a square lattice or network. These elements respond non-linearly to their current state of stress. When stress exceeds a threshold value, elements discharge energy to connected elements. Element stresses accumulate from a global driving function as well as from discharges from neighboring elements. This simple model can be intuitively interpreted to describe the behavior of a sand pile on a finite table. The sand pile model is quite simple and can be used as a metaphor for many systems besides the original example. One starts with a set of boxes on a regular lattice into which one begins dropping grains of sand at random. When the number of grains in a particular box exceeds a critical value, it spills its sand into neighboring boxes. If neighboring boxes are near the critical value, they can be pushed over and also spill. When sand reaches the edge of the lattice, it falls off and out of the system. Incredibly, this very simple set of rules, operating within a regular lattice, generates complex behavior (see **Figure 4**). The system self-organizes into a critical state where events of all sizes can occur at any time and thus are, in some sense, unpredictable.

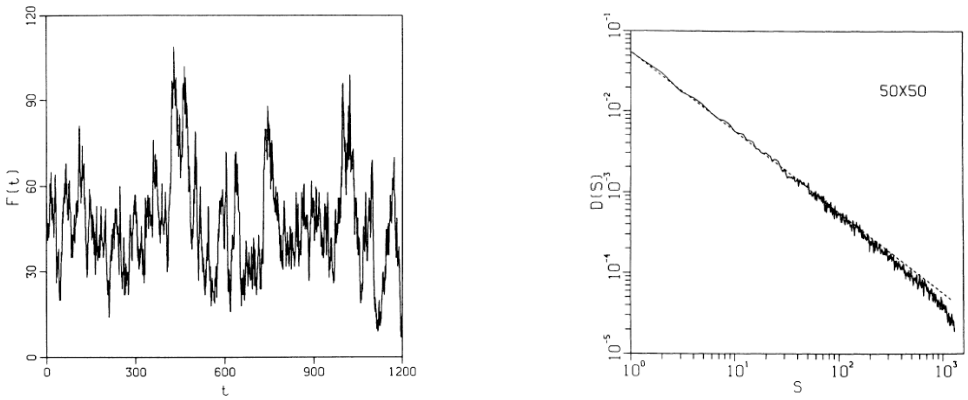


Figure 4. Example time series and frequency distribution of cascade size in the sandpile model. From Bak, Tang and Wiesenfeld [1988].

Since being introduced by BTW, this simple model and the concept of SOC caught on like wildfire and created its own avalanche of activity in a wide range of fields [see Buchanan, 2001]. There have been over 1800 citations of BTW's paper within the scientific literature. The sand pile metaphor and accompanying model has been used with slight modifications to describe phenomena as seemingly disparate as earthquakes, cascading failures of power grids, price fluctuations on the stock market, and the spread of infectious diseases. In all these interpretations, a node has only two states. For a physical node (block of rock, computer, relay, etc.), the node is either slipped/stationary, on/off, tripped/untripped, etc. For a human node, the

state represents a binary decision, yes/no, act/acquiesce, buy/sell, or a state such as healthy/sick. When one node changes state, it influences the states of its neighbors, i.e., it increases their stress, sends current their way, influences decisions, infects them, etc. All of these simple computational models yield macroscopic SOC behavior characterized by temporal-spatial fractals. This macroscopic behavior is also seen in the corresponding natural or engineered system. For such behavior, no single characteristic event size exists, nor is there a characteristic spatial or temporal scale. The simplifying aspect is that power laws describe the statistical properties, i.e., they are fractal. To obtain full generality, the exponent for the power law distribution would be identical for systems that appear to be different from a microscopic perspective. Attempting to establish this relationship is an active avenue of research for physicists.

Physicists work to simplify, so as to distill the essence of what is causing certain macroscopic behavior. BTW's sand pile model and its daughters exemplify this approach. In general, most advocates of SOC contend that the details underlying whether a node is in one state or another don't matter. What matters, is that the ultimate behavior of a node is discrete and it that it influences the state of its neighbors when it changes state. Despite their sometimes outrageous simplicity, these models have been found to reproduce important statistical properties of real systems (especially from the standpoint of extreme event frequency). However, an outstanding question in many cases is whether they produce that behavior *for the same reasons* that the real systems do, thereby helping us understand these systems, or whether the agreement is circumstantial [e.g., see the recent paper of Willinger et al., 2002]. While such modeling, with very wide potential applications, has generated a great deal of rhetoric, the models have not been systematically explored and organized. A clear-cut and generally accepted definition of SOC does not exist, nor are the necessary conditions under which SOC behavior arises fully understood. There is also a lack of mathematical formalism and framework.

Finally, we note that recently it has been found that SOC is not the only explanation for power-law spectra. Highly Optimized Tolerance or HOT (Carlson and Doyle, 1999) is a mechanism that produces power-law spectra through optimization of system response to random perturbations. In SOC systems, complex behavior arises from state changes among interacting elements, but the patterns of interconnection and the rules governing state transitions are fixed. In contrast, HOT systems can change their structural properties over time. This is a fundamentally different route to complex behavior. Infrastructures, like other complex systems, change over time, and understanding the factors that drive and constrain these changes is crucial if we seek to improve infrastructure performance and reliability.

We've learned that many systems of interacting simple components can undergo widespread state changes or cascades. The propensity to cascade depends on the system structure, rules the components follow, the way they influence each other, and how the system is driven. From the BTW paper and successors, we have seen that cascading behavior can arise when the component rules, network connections, and driving process are fixed within certain ranges. We also have the suggestion that many kinds of systems (plate boundaries, forests, etc.) may be driven to exhibit cascade behavior, perhaps by accumulating energy until a critical state is achieved among the components. In this sense the system might naturally converge to a critical state. This interesting cascade behavior and the provocative suggestion of ubiquitous criticality only depend on the dynamics of component state transitions; the transition rules, parameters (such as transition thresholds), and connections are all fixed. While the assumption of fixed network structure and fixed transition/propagation rules may be true of some systems and over some time scales, it is not true of infrastructures over long time scales.

2.1.3 Growth, Evolution and Adaptation: Complex Adaptive Systems

In most real systems, the rules, parameters, and connections that make up a network change over time. How might the slower dynamics that can cause component behavior, parameters, and connections to change over time influence cascades or other system-level phenomena? Can the plasticity of component behavior and network structure make the system more robust, or less robust, or provide robustness to some perturbations at the expense of increased vulnerability to others? What things might drive changes in component behavior and network structure? Is there some global performance objective (e.g. an airline's planning of its routes) or do changes result from local optimization decisions (e.g. locating a freeway off-ramp)? Are there adaptive pressures on system properties (e.g. susceptibility to cascading) that tend to shape component behavior or network geometry even if they are not explicit design goals?

Answers to such questions are important for understanding how complex systems such as infrastructures behave over extended periods, and what effects policies might have on their future shape and performance. Such insights could help design policies leading to more robust systems. Conversely, neglecting long-term growth processes may lead to ineffectual strategies: reinforcing "weak links" in an existing network is not helpful if the processes driving network growth cause the creation of such features.

A network's current structure and dynamics together determine its performance. These elements change over time, and those changes may depend in some way on the system's overall performance. **Figure 5** conceptually illustrates this feedback process, in which either local or global performance measures guide the modification of the local properties of the system. To understand this larger process, we must understand which elements of the network (e.g. structure, transition rules, propagation rules) are subject to change, and how the performance of the system influences these changes.

This problem is not unique to infrastructures. Many other complex systems (e.g. ecological systems, economic systems, metabolic systems) are characterized by gradual change in the parameters that determine their short-term dynamics. These changes often arise from perturbations in "local" features (such as the rules determining the behavior of individual elements) that produce changes in global system properties. If the values of these global properties feed back to the perturbation process, the system can adapt over time to the external stresses reflected by this feedback.

This structure, called a Complex Adaptive System or CAS, has been used to model a wide range of natural and engineered systems. The CAS formalism is widely applicable because it is very general: it posits a set of interacting components whose actions and interactions are shaped, over time, by (often emergent) system-level properties. Applying the formalism requires specifying the ranges of actions and interactions (as with any complex system), and identifying the feedback mechanisms that relate system performance to changes in those elements. The challenge for infrastructure modeling is to discover these mechanisms.

Other approaches also consider feedback from system performance to structure. Highly Optimized Tolerance (HOT) is a mechanism proposed by Carlson and Doyle [1999] to account for the widespread observation of power-law scaling in both natural and designed systems. Their explanation makes use of the fact that the rules and parameters governing the propagation of stresses through a system are themselves subject to change over time, and that these changes can be shaped by intention or by adaptive pressures to optimize system performance with respect to some objective. The objective they consider is a loss function formed from the product of the probability that an event leading to a disruption will occur and the cost of the disruption. The cost is assumed to scale as a power of the "size" of the event, for example the number of nodes

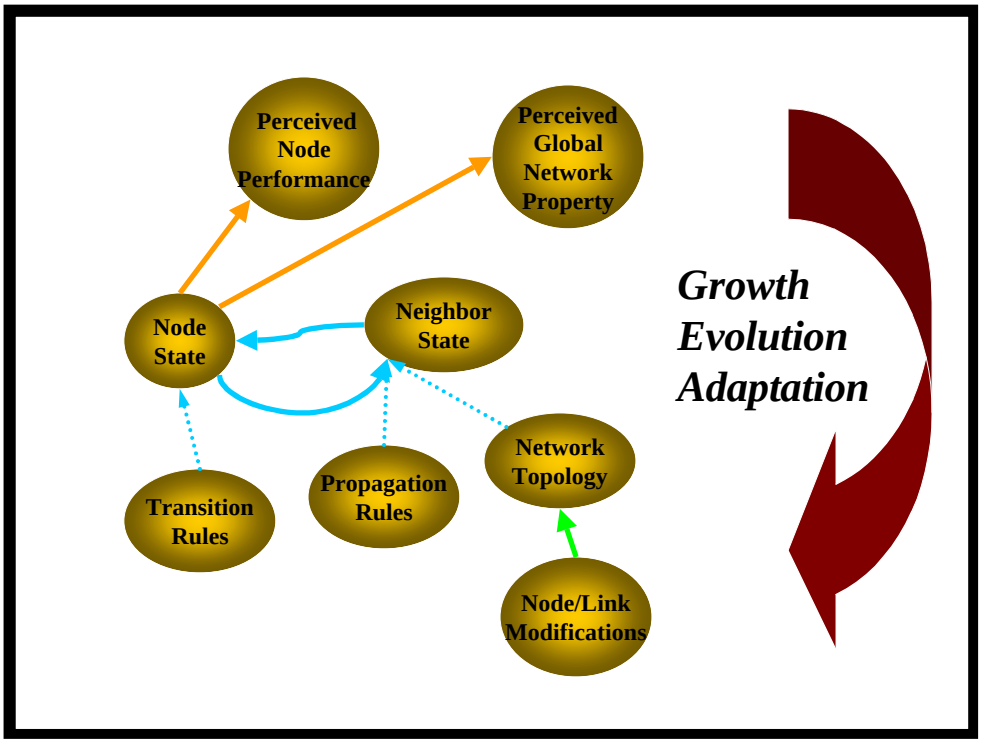


Figure 5. Elements and interactions in a complex adaptive network.

that experience a state change as the disruption moves through the system. They hypothesize that, either through design or adaptation, the system can redistribute a generic resource in order to control the event size. The resource is conserved, so that decreasing the "size" of one event entails increasing the "size" of others. When combined with plausible assumptions about the probability distribution of initiating events, this very general formulation yields a power-law relationship between event size and probability. HOT systems differ from SOC systems in two important respects. First, by deploying resources to limit the extent of disruptions, they can maintain states that would be far beyond the cascade threshold for SOC systems. Second, because they deploy resources in response to a specific probability distribution for initiating events, their performance can be greatly degraded if this distribution changes: optimization for one set of conditions creates vulnerabilities to changes in those conditions.

The HOT mechanism accounts for the effect of a global objective function on the configuration of states and the consequent distribution of cascades, but does not consider changes in connections or behavior. Resource deployment in HOT is shaped by a global objective function, but system changes in natural and engineered systems may also be shaped by local optimization. For example, the scale-free network naturally "evolves" from a local decision to preferentially connect to the well connected Such networks turn out to be robust against random failures, but the relationship, if any, between the local rule and the global property is unclear. It would be interesting to know if the rule is an adaptive response to constraints that favor the global property. In general, it seems that an evolutionary perspective will be helpful for understanding infrastructure networks because they tend to be modified incrementally, and the increments must tend to enhance the performance of the system in some way.

2.2 Summary of Model Technologies

Nonlinear dynamical system models are usually difficult or impossible to solve analytically. Widespread availability of computers has enabled the recent advances in understanding such systems. Models of complex systems are all computational. Very few relevant results can be obtained through analytic methods where simplifying assumptions are required to make progress. Complex system models can be generally lumped into a single overarching class: Agent Based Models (ABM) of which standard percolation, Ising, Sand-pile, and cellular automata (CA) form a progression of increasing complication (or degrees of freedom). There is also a progressive complication of agent (node) degree or interaction network topology which can be incorporated within ABM: regular lattices, (N-1) dimensional networks, fractal networks, and finally time-varying connections.

ABMs have the ability to span the spectrum from very simple to quite complicated, especially when applied to disparate fields such as Sociology, Anthropology, Economics, Engineering, and Biology/Ecology. In these kinds of applications, ABM show great potential in being able to capture 'real' system behavior because they:

- Allow heterogeneity within agent types;
- Allow multiple agent types;
- Allow adaptation of agent behavior;
- Allow dynamic changes in connections;
- Are open ended with respect to complication and the questions that one may ask;
- Are calculational and computers are now fast and plentiful.

A number of models used to study adaptive systems can be characterized as "artificial chemistries." These models have been applied to solve optimization problems, to study the origins of pre-biological replication, to examine the creation and persistence of specialized organizations, and for many other purposes. Artificial chemistries (ACs) are analogous to ABM but use chemistry as a metaphor for describing the elements and processes of the system being studied. AC is a potentially powerful formalism for studying the long-term behavior of interconnected infrastructure systems. While these processes can be modeled in several ways, AC models provide an intuitive description for the flow and transformation of materials in systems that also generate new materials and material uses. Like the other formulations we have explored for describing complex systems, AC models can be understood as specializations of agent models. There are ways agent models can be used to realize an artificial chemistry. For example, we can define three agent sub-types corresponding to molecules, reactions, and reaction sites. Network structures can be used to characterize the accessibility of reaction sites to reactants.

2.3 Infrastructures as Complex Adaptive Systems

Infrastructures are systems that deliver basic services, such as power, water, food, and communications to widely distributed users. Economies of scale and spatial variations in resource availability cause these services to be produced in specialized large-scale facilities. Infrastructures are consequently composed of localized production facilities, distributed points of consumption, and transport systems that connect these locations. Networks are a natural and commonly used way of conceiving of infrastructure systems. The components of an infrastructure naturally fall into local elements, representing sites of production, consumption, or transfer, and connecting elements that exchange materials or control information. These elements are readily associated with the nodes and connections of a network.

Although infrastructures are sometimes described using simple network models having nodes and connections with uniform properties, this representation excludes potentially important information. The components of real networks are usually heterogeneous. Although individual network components do not often have the same properties, they can be lumped into similar classes of units that have similar behavior. There are often many such classes, and individual units are often heterogeneous within their class. Although informative models have been constructed assuming all network connections have the same properties, they are usually also heterogeneous in real systems. Connections could have different 'strengths' (e.g. capacities) and 'types' (e.g., unidirectional vs. dual directional), and different nodes may have differing degrees.

The existence of heterogeneities in infrastructures can be easily accommodated in network models, however little theoretical work has evidently been done in this area. It is currently unclear, for example, how the critical parameter values for the onset of cascade failure depend on heterogeneities. This problem has special salience for our work. While it is appealing to describe interdependent infrastructures as a single network, the composite network is clearly heterogeneous, with nodes representing very diverse types of facilities (e.g. power plants, natural gas collection stations, refineries, etc.) and connections representing the flows of materials, information, and the interconnections between infrastructures.

The physical elements of infrastructures are not the only source of interdependencies. In order to perform their functions, infrastructures require people as operators, decision makers, policy makers, and in other roles. The people and organizations that use, operate, and shape infrastructures can be understood as *social networks* of various kinds. These social networks interpenetrate the physical network of infrastructure hardware, and a single social network might interact with more than one physical network. Interdependencies also arise from infrastructure growth. The topology of one infrastructure network may constrain the way another expands, for example because of shared rights-of-way.

Thus, infrastructures have many, and sometimes all, of the underlying characteristics of the complex systems studied by the physicists, biologists, sociologists, anthropologists, economists. There is a growing body of data and analysis documenting that infrastructures are composed of complex networks, show cascading and power law behavior for event size frequency distributions, and are in essence, Complex Adaptive Systems. It is possible that some infrastructures (or the composite interdependent infrastructure amalgamate) will have regimes within parameter space where the system will converge naturally to a critical state, or in other words, exhibit SOC. There, they will be intrinsically unpredictable. However, it is also possible that infrastructures will show such behavior not because they are in an SOC state, but because they are in a HOT state because, at least at some level, all individual infrastructures have been 'designed' or have 'evolved' subject to constraints and performance criteria.

While the state of understanding of CAS suggests that it is clearly applicable to interdependent infrastructures, there are also important gaps in this understanding. We discuss these gaps in the next Section. Section 4 outlines our approach to closing them.

3. Basic Research Issues

Review of the complexity literature indicates that there are common unresolved issues pertaining to the application of complex adaptive modeling to real-world situations. We have distilled these into 5 main categories:

- Model and network abstraction (Section 3.1)
- Model and network behavioral definition and categorization (Section 3.2)
- Hazard/problem and solution/fix categorization (Section 3.3)
- Adaption/evolution/directed-evolution (Section 3.4)
- Predictability (Section 3.5)

While there is a wide range of issues that can be raised, these are a starting point. As research progresses, new issues will undoubtedly be identified, and the relative importance of existing research issues may change. The five basic research issues discussed here are not independent of one another. Model and network abstraction, model validation, temporal and spatial scaling, feedback loop structure, and network topology pervade all aspects of network and infrastructure analysis.

3.1 Model and Network Abstraction

The model and network abstraction issue can be summarized as the need for research to determine the appropriate level of simplification required to produce an accurate model, and establish the optimal metrics for model evaluation. Computational models require the abstraction of complex processes into comparatively simple algorithms in order to form tractable simulations. Models containing complex and detailed algorithms may appear to closely simulate complex realities; however, for very detailed models, parameters can become interdependent and the utility of the model as a predictive tool is thereby diminished. Alternatively, for less detailed models, complex real phenomena are reduced to very simple algorithms. These models may be more useful predictive tools and facilitate real-time analyses, but they may lack adequate representation of true interdependencies and event consequences. The goal of any research effort is to find the optimal level of simplification appropriate to a given problem so that the essential processes are accurately represented, and excessive detail is discarded. Simplification to the most essential processes allows greater understanding of the essential nature of the driving forces relevant to a problem or situation.

Model validation provides the best means of determining if a computational model includes an optimal level of detail; however, the metrics chosen to validate a computation model can determine the level of success achieved by the model. Full model validation must comprise:

- Simulation of micro and macro behaviors.
- Simulation of phenomena at multiple time scales.
- Accommodation of imperfect data streams.
- Parametric studies to provide model system behavior through the full range of possible parameter space.
- Prudent interpretation of results, especially when results are non-intuitive.

Since some model validation metrics may favor specific models, the metrics used for evaluation must be carefully considered.

3.2 Model and Network Behavioral Definition and Categorization

The study of network behavior has come to employ a taxonomy of network geometries (such as random, regular, small-world, and scale free). Separately, studies of cascade behavior have posited simple behavioral rules for nodes, typically using a regular network geometry, and identified relationships between parameter values for the behavioral model and the extent of disturbance. Some recent analyses (e.g. Sachtjen 2000) have looked at the effect of network topology on cascade behavior given a specific behavioral model. Current understanding of how network topology and the behavior of nodes and connections interact to create distinct regimes of macro behavior appears to be fragmentary. There is an evident need for systematic exploration and parametric studies to map out the ranges of network topology and node behavior, and how macro behavior changes across this space of possibilities.

A systematic understanding of how topology, micro behavior, and macro response interact would also help focus our investigation of broader problems. Network structure and behavior are not fixed. Networks grow, evolve, and adapt based on the network topology, rules for node interaction, and imposed constraints, which may be implemented on different spatial and temporal scales. Network adaptation, discussed below in Section 3.4, might be understood as a trajectory through the space of topologies and behaviors. Understanding the properties of this "landscape" would be a great help in understanding how systems move through it over time, and what features they may encounter. Different network topologies, behavior rules, and adaptation schemes may be a very general framework for categorizing and understanding infrastructures.

3.3 Hazard/Problem and Solution/Fix Categorization

Section 3.2 discussed the use of network topology and adaptive behaviors as methods of characterizing infrastructures for more efficient analysis. In this section we discuss the utility of categorizing infrastructures according to their vulnerabilities, or alternatively, according to the methods of vulnerability mitigation.

Categorization by hazard or problem, suggests that the characteristics of system failure, or the characteristics of the resulting hazard may provide the best criteria by which to group systems for analysis. For example, an underlying feature of the network geometry may control how, and how far, events propagate. Metrics might include the event size, probability, and post-event network conditions such as connectedness and path length.

Alternatively, the method by which a system is made more stable may provide an efficacious categorization criterion. This is referred to here as solution/fix categorization. Some examples of categories of solutions include:

- Enhance directed response to quickly eliminate any instability in the system. Improve early detection. Note that condition-dependent response might be modeled as changes to local node rules, or might include changing node or connection behavior based on the states of connected nodes. The latter kind of response entails a monitoring and communication system operated in conjunction with the original network, and therefore introduces an important general interdependency. Understanding the performance of this kind of response system as a function of general parameters, such as the relative time constants between the original network and the signal/control network and the geometries of the two networks, is especially relevant for NISAC.
- Add security or strength to existing system. This might be modeled as a general or localized decrease in the transition probability of nodes.

- Add redundancy to critical links in system. Understanding how such links are selected and reinforced may also provide relevant information for modeling network adaptation.
- Add buffers to keep system perturbations from causing system failure. As with link reinforcement, understanding which modifications are effective and how modifications are made in real systems will give insight for formulating adaptation rules.
- Enhance general response – e.g., communicate hazards to cause best public response. Such responses may also require a representation of communication and control systems.

3.4 Adaption/Evolution/Directed-Evolution

System behavior and system structure are linked, and evolve through adaptive feedback. The choice of the feedback loop is therefore essential to an accurate simulation. There are many research issues associated with the appropriate choice of feedback loop. The variable chosen for the feedback signal, the function selected for optimization, and the structure of the feedback loop itself, represent a significant research areas. One possible scenario for the complex interplay of these factors was shown in **Figure 5**.

Economic signals form much of the feedback influencing infrastructures. Choices of economic objective functions have a strong influence on infrastructure development and operation. Understanding these interactions can help us answer questions such as:

- Are JIT systems brittle or flexible? Do they provide greater or lesser resilience than systems with large buffers?
- What happens in an economy based on sustainability rather than growth?
- How will infrastructures change in the next 10 –20 years? How do market structures influence that evolution?

Not only is the feedback loop important, but also the temporal and spatial scales over which the feedback process is examined can be critical. For example, averaging over large temporal and spatial scales could cause critical signals to be diffused. Inversely, many real systems do not react to phenomena that occur at high frequencies. Spatial scales must also be chosen appropriately since local optimization decisions may have adverse effects on a global performance objective, and vice-versa.

3.5 Predictability

Complexity science pertains to phenomena that can be described on a macro scale by physical laws, yet prediction of precise occurrences and magnitudes of these phenomena remains elusive. This research issue explores the distinction between predictable and unpredictable phenomena, and in particular, the policies and controls that can be applied to unpredictable phenomena. We would ideally map the boundaries between predictable and unpredictable conditions, and controllable and uncontrollable systems, as a function of general system parameters. Again, an understanding of the “landscape” of network topology and behavior is required.

The research questions under the umbrella of predictability include:

- Is there an identifier that indicates whether an event is predictable or unpredictable?
- Is there a root cause of predictability or unpredictability? Such causes, expressed as critical parameter values, are found in simple chaotic systems. Can we find similar results for complex networks?

- If a phenomenon is not specifically predictable, are there mechanisms that can be used to control some aspect of the phenomenon, such as its magnitude, time of occurrence, or severity? Rather than controlling an unpredictable phenomenon, can we control the system in a way to make it (more) predictable? If so, what are the consequent costs or inefficiencies under normal conditions?
- What are the optimal methods for communications concerning unpredictable phenomena?
- How do we assess model validity, or judge the confidence we should place in model results, if the modeled system is unpredictable?

Parametric studies predict how an end result or macro behavior may change depending upon input properties or micro behaviors. These studies are useful for assessing the effects of policy changes, and locating transitions between different domains of macro behavior.

4. Definition of the R&D effort

The work of the Infrastructure Complexity R&D Group is one of NISAC's long-term investments in understanding infrastructures and their interdependencies. Our mission is to identify theories, methods, and analytical tools from the study of complex systems so that they can be brought to bear on critical infrastructure problems. This year's efforts were conducted according to the following charter:

> *Conduct research on complex system theory and modeling to identify potentially useful, new or additional modeling, simulation and analysis methodologies and tools. Begin to develop and test new models, simulation and analysis tools. Evaluate their utility. Continue evaluation of models and data related to infrastructure interdependencies and system evolution. Conduct research to anticipate evolutionary trajectories of infrastructures and their interactions. Identify potential limitations and vulnerabilities arising from patterns of infrastructure evolution to allow development of more robust systems.*

This year's work in Complex Systems led to a shared understanding of current models used in the field, an identification of those approaches with obvious relevance for infrastructures and interdependencies, and insights into which actions we can recommend to expand and apply current complex systems theory. We believe the charter as stated above continues to reflect our goals and activities into the foreseeable future. However, we will revisit this charter at the end of each year to insure that our work provides the needed support to NISAC.

In this past year, we have developed insights into three areas: how we can conduct this research to most effectively develop and apply analytical insights; which technical questions most urgently need answering; and which formulations show the most promise for providing answers. Below we present our recommendations in each of these three areas designated by Process (Section 4.1), Technical Objectives (Section 4.2) and Modeling Tools (Section 4.3).

4.1 Process

Our research process is meant to make the best insights and techniques available for the infrastructure problems NISAC will be tasked to analyze. Meeting this goal entails gathering relevant information where it exists, advancing the current understanding in those areas where we need it most, and making the best approaches available for application. This year we have

reviewed diverse papers on current theory and application in Complex Systems. This year's review laid essential groundwork, but assimilating the literature will continue to be important. Complex systems research is very active and is being applied in many areas. It is important to track advances so that our work remains focused on unanswered questions, and so that others' insights can be quickly applied to our problems.

Continued information collection is critical for identifying promising modeling methods. Disseminating that information to model developers and users is equally important. We have used regular open discussions to share ideas among NISAC technical staff, and believe it is important to sustain and expand this communication. We propose to continue these discussions on a regular basis, and will seek to involve interested members of other groups working in complex systems. We can also use these discussions to help generalize the specific detailed work of NISAC so that it can be more easily and quickly applied to new problems. Expanding the pool of participants to include non-NISAC researchers will provide additional insights into infrastructure behavior and modeling.

We must encourage the commitment of participants and customers if the research program is to yield improvements to methods and analyses. Sustaining commitment is difficult because of the numerous short-term demands that inevitably arise to support NISAC's growing customer base. In order to keep participants and customers engaged, we propose to regularly document specific research goals, activities, and results. The format and frequency of these documents will be tailored to the specific findings we need to communicate, but will include journal articles, SAND reports, and conference papers. Regular documentation of research results will help focus participants' attention on continued progress, and will keep customers regularly informed about our products and direction. We cannot anticipate the results of research, and so cannot fully specify the series of documents we will produce. We can, however, set short-term goals that we will pursue in the immediate future. After these goals have been accomplished, we will define more specific endpoints for the next phase of our work. The content and format for the first phase of next year's work is discussed further in the section on Technical Objectives below.

4.2 Technical Objectives

This year's research has identified several technical issues, detailed in Section 3, whose resolution will identify how complex systems theory can best inform the modeling of infrastructures and their interdependencies. Our general technical objectives are designed to investigate these issues, to identify the general characteristics of infrastructures that govern their behavior and interactions, and to propose or develop models that can best support NISAC's analyses. Next year's research will make progress in all of these areas.

For the first phase of next year's work, we have defined a limited set of specific short-term objectives, which we intend to pursue by incrementally expanding and analyzing the simple demonstration model we have developed this year (see Appendix C). The results of this investigation will be documented in a journal article. The goals for this work are to:

- Fill in gaps in the current understanding of how simple transition rules and network geometries can combine to permit system-spanning cascades;
- Discover the ranges of parameter values that cause transitions in macro-scale behavior (e.g. between isolated failure and pervasive failure);
- Identify the influence on cascading of departures from common assumptions in current models. Some specific variations will include diffusive relaxation of node "charge", heterogeneity in node and network properties, etc.

In the second phase of next year's work, we will begin to analyze other outstanding issues. At this time we cannot define the order in which these problems will be approached, nor the specific procedures we will use, however we expect that the simulations done in the first phase might be extended to consider some subset of these problems. We will place a high priority on examining the effects of interdependencies on system behavior. In addition to looking at new system features (such as multiple interdependent networks), we will also use simulation studies as a tool to explore some of our general modeling questions. For example, predictability may be assessed through parameter uncertainty analyses, and the differences we observe between more specific and less specific representations of the same system may provide general guidelines for judging the appropriate model resolution for specific problems.

Some possible simulation studies, which may be the focus of subsequent phases, are outlined below:

- Focus on interdependency by considering a social network connected to a transportation network on which material and people move, coupled with a power network and a telecommunications network. Within this context, consider some disruption and the flare up of civil disorder. The analysis would imply using ABM with agent connectivity defined differently within each sub-network or infrastructure (some regular, some scale-free, some random or small world).

- Examine the spread of infectious disease in livestock under various detection and mitigation scenarios. This process involves the interaction of transportation, communication, and government service infrastructures, each characterized by distinctive time constants and connectivity. The analysis might focus on policy evaluation under different objective functions.

- Use artificial chemistry (AC) as a formalism for studying the long-term behavior of interconnected infrastructure systems. AC models provide a natural set of metaphors for describing the flow and transformation of materials in systems that also generate new materials and material uses.

4.3 Modeling Tools

The models of complex systems that are discussed in the current literature can be readily recast as agent-based models (ABM). Agents often are used to model very specific and complicated entities (such as power plants), but can also implement the simple transition rules typical of published complexity studies (such as having a binary state variable representing working/failed, uninfected/infected, etc.). Agents update their state based on interactions with one another and with their environment. All agents in a model might belong to a single agent type, or a model might include agents of different kinds. The topology of agent interactions can be described using the kinds of networks currently discussed in the literature. Agent modeling tools such as RePast allow us to readily reproduce these simple formulations, and to easily extend them. For example, interactions can be governed by relative spatial location or by networks derived from data on actual infrastructures. Because of the flexibility of agent models, and the ready availability of agent modeling tools, we anticipate that we will use this technology to formulate and answer many of our specific research questions.

In comparison to other agent modeling systems currently being developed and used within NISAC, our models will have a high degree of abstraction. This will allow us to better understand how general features, such as network connectivity and threshold size, can control

system behavior. Such a general understanding may allow us to find model simplifications and mitigation strategies with application to many specific systems.

5. References

Barabasi, A.-L., and R. Albert, Emergence of scaling in random networks, Science, 286:5439:509-512, 1999.

Albert, R., and A.-L. Barabasi, Statistical mechanics of complex networks, Reviews of Modern Physics, 74:1:47-97, 2002.

Albert, R., H. Jeong, A.-L. Barabasi, Error and attack tolerance of complex networks, Nature, 406:378-382, 2000.

Bak, P., *How Nature Works: The science of self-organized criticality*, Copernicus, Springer-Verlag New York, Inc., 1996.

Bak, P., C. Tang, and K. Wiesenfeld, Self-Organized Criticality, Physical Review A, 38:1:364-374, 1988.

Bak, P., C. Tang, and K. Wiesenfeld, Self-Organized Criticality: An explanation of 1/f noise, Physical Review Letters, 59:4:381-384, 1987.

Binney, J.J., N.J. Dowrick, A.J. Fisher, and M.E.J. Newman, *The Theory of Critical Phenomena*, Oxford, UK: Oxford University Press, 1992.

Buchanan, M., *Ubiquity: the science of history… or why the world is simpler than we think*, Crown Publishers: New York, 2001.

Carlson, J.M, and J. Doyle, Highly optimized tolerance: A mechanism for power laws in designed systems, Physical Review E, 60:2:1412-1427, 1999.

Dorogovtsev, S.N., and J.F.F. Mendes, Evolution of Networks, Advances in Physics, 51:4:1079-1187, 2002.

Erdos, P., and A. Renyi, On Random Graphs, I, Publicationes Mathematicae (Debrecen), 6:290-297, 1959.

Gleick, J, *Chaos: Making of a new science*, Viking: New York, 1987.

Jensen, H.J., *Self-Organized Criticality: Emergent complex behavior in physical and biological systems*, Cambridge lecture notes in Physics, Cambridge University Press, 1998.

Lorenz, E. N., Deterministic nonperiodic flow. J. Atmos. Sci. Vol. 20, pp 130-141, 1963.

Milgram, S., The small world problem, Psychology Today, 2:60-67, 1967.

Nicolis, G., 'Physics of far-from-equilibrium systems and self-organization', in P. Davids (ed.), *The New Physics*, Cambridge University Press, 1989.

Sakhtjen, M.L., B.A. Carrerras and V.E. Lynch, Disturbances in a power transmission system, Physical Review E, 61:5:4877-4882, 2000.

Stauffer, D., and A. Aharony, *Introduction to Percolation Theory*, Second Edition, Taylor & Francis: London, 1992.

Strogatz, S.H., Exploring Complex networks, Nature, 410:268-276, 2001.

Waldrop, M.M., *Complexity: The emerging science at the edge of order and cha*os, New York: Penguin, 1992.

Watts, D.J., and S.H. Strogatz, Collective dynamics of 'small world' networks, Nature, 393:440-442, 1998.

Willinger, W., R. Govindan, S. Jamin, V. Paxon, and S. Shenker, Scaling phenomena in the Internet: Critically examining criticality, PNAS, 99:suppl. 1:2573-2580, 2002. www.pnas.org/cgi/doi/10.1073/pnas.012583099.

Appendix A: Our Course of Study

A large component of our R&D effort this year was focused on an examination of the scientific literature in Complexity Science. This literature is extremely diverse and is rapidly encompassing almost every field of scientific investigation. The sorting of this literature to first identify the critical set of topics, and then to find the critical papers in each, took a significant effort (and made all our heads spin from time to time!). We covered these critical papers in a series of Sessions where participants would first read the assigned papers and then come together to discuss them in a 2 hour period. Below we list these readings by session and topic. We also include some 'additional reading' that might also be of benefit to fill out the topic a bit more. Note that a CD that contains copies of all the research papers listed below is available from the Infrastructure Complexity R&D Group. Books must be purchased separately.

Preliminary: Background and Overview

Buchanan, M., *Ubiquity: the science of history… or why the world is simpler than we think*, Crown Publishers: New York, 2001.

Additional Reading and Reference:

Jensen, H.J., *Self-Organized Criticality: Emergent complex behavior in physical and biological systems*, Cambridge lecture notes in Physics, Cambridge University Press, 1998.

Waldrop, M.M., *Complexity: The emerging science at the edge of order and cha*os, New York: Penguin, 1992.

Bak, P., *How Nature Works: The science of self-organized criticality*, Copernicus, Springer-Verlag New York, Inc., 1996.

Bar-Yam, Y., *Dynamics of Complex Systems*, Studies in Nonlinearity, The Advanced Book Program, Perseus Books: Reading, Massachusetts, 1997.

Stauffer, D., and A. Aharony, *Introduction to Percolation Theory*, Second Edition, Taylor & Francis: London, 1992.

Gleick, J, *Chaos: Making of a new science*, Viking: New York, 1987.

Lorenz, E. N., Deterministic nonperiodic flow. J. Atmos. Sci. Vol. 20, pp 130-141, 1963.

Nicolis, G., 'Physics of far-from-equilibrium systems and self-organization', in P. Davids (ed.), *The New Physics*, Cambridge University Press, 1989.

Binney, J.J., N.J. Dowrick, A.J. Fisher, and M.E.J. Newman, *The Theory of Critical Phenomena*, Oxford, UK: Oxford University Press, 1992.

Session 1: Self-Organized Criticality… From Physical to Social Sciences

Bak, P., C. Tang, and K. Wiesenfeld, Self-Organized Criticality: An explanation of 1/f noise, Physical Review Letters, 59:4:381-384, 1987.

Sakhtjen, M.L., B.A. Carrerras and V.E. Lynch, Disturbances in a power transmission system, Physical Review E, 61:5:4877-4882, 2000.

Watts, D.J., A simple model of global cascades on random networks, Proc. Of the National Academy of Sciences, 99(9); 5766-5771; 2002, www.panas.org/cgi/doi/10.1073/pnas.082090499.

Additional Reading:

Bak, P., C. Tang, and K. Wiesenfeld, Self-Organized Criticality, Physical Review A, 38:1:364-374, 1988.

Session 2: Beyond SOC… and what about model validation?

Carlson, J.M, and J. Doyle, Highly optimized tolerance: A mechanism for power laws in designed systems, Physical Review E, 60:2:1412-1427, 1999.

Willinger, W., R. Govindan, S. Jamin, V. Paxon, and S. Shenker, Scaling phenomena in the Internet: Critically examining criticality, PNAS, 99:suppl. 1:2573-2580, 2002. www.pnas.org/cgi/doi/10.1073/pnas.012583099.

Session 3: Introduction to Complex Networks

Albert, R., H. Jeong, A.-L. Barabasi, Error and attack tolerance of complex networks, Nature, 406:378-382, 2000.

Strogatz, S.H., Exploring Complex networks, Nature, 410:268-276, 2001.

Additional reading:

Buchanan, M. Nexus: Small Worlds and the Groundbreaking Science of Networks. W. W. Norton & Company Inc., NY NY, 2002.

Callaway, D.S., M.E.J. Newman, S.H. Stogatz, and D.J. Watts, Network Robustness and Fragility: Percolation on Random Graphs, Physical Review Letters, 85:25:5468-5471, 2000.

Session 4: Adaptive Networks

Feitelson, E., and I. Salomon, The implications of differential network flexibility for spatial structures, Transportation Research Part A, 34:459-479, 2000.

Dorogovtsev, S.N., and J.F.F. Mendes, Evolution of Networks, Advances in Physics, 51:4:1079-1187, 2002.

Sessions 5, 6 and 7: In-depth Graph Theory as relevant to Complex Networks

Albert, R., and A.-L. Barabasi, Statistical mechanics of complex networks, Reviews of Modern Physics, 74:1:47-97, 2002.

Additional Reading:

Watts, D.J., Small Worlds: *The dynamics of Networks between Order and Randomness, Princeton Studies in Complexity*, Princeton University Press: Princeton, New Jersey, 1999.

Barabasi, A.-L., *Linked: The new science of networks*, Perseus Publishing: Cambridge Massachusetts, 2002.

Session 8: Starting to dig into Social Behavior and Agent-Based Models

Axtell, R., Why Agents? On the varied motivations for agent computing in the social sciences, Center on Social and Economic Dynamics, The Brookings Institution, Working Paper #17, 2000.

Berry, B.J.L., L.D. Kiel, and E. Elliot, Adaptive agents, intelligence, and emergent organization: Capturing complexity through agent-based modeling, PNAS, 99:suppl. 3:7187-7188, 2002. www.pnas.org/cgi/doi/10.1073/pnas.092078899. (Overall introduction to special PNAS issue of the same name)

Additional Reading:

Epstein, J.M., and R. Axtell, *Growing Artificial Societies: Social Science from the bottoms up*, Brookings Institution Press: Washington, D.C., 1996.

Session 9: Modeling Social Behavior

Epstein, J.M., Modeling civil violence: An agent-based computational approach, PNAS, 99:suppl. 3:7243-7250, 2002. www.pnas.org/cgi/doi/10.1073/pnas.092080199.

Danielson, P., Competition among cooperators: Altruism and reciprocity, PNAS, 99:suppl. 3:7237-7242, 2002. www.pnas.org/cgi/doi/10.1073/pnas.082079899.

Additional reading:

Axelrod, R., *The complexity of cooperation: Agent-based models of competition and collaboration*, Princeton Studies in Complexity, Princeton University Press: Princeton, New Jersey, 1997.

Session 10: Modeling Financial Behavior

Tesfatsion, L., Economic agents and markets as emergent phenomena, PNAS, 99:suppl. 3:7191-7192, 2002. www.pnas.org/cgi/doi/10.1073/pnas.072079199.

LeBaron, B., Short-memory traders and their impact on group learning in financial markets, PNAS, 99:suppl. 3:7201-7206, 2002. www.pnas.org/cgi/doi/10.1073/pnas.072079699.

Additional Reading:

Eguiluz, V.M., and M.G. Zimmermann, Transmission of information and herd behavior: An application to financial markets, Physical Review Letters, 85:26:5659-5662, 2000.

Session 11: Modeling Policy

Lempert, R., Agent-based modeling as organizational and public policy simulators, PNAS, 99:suppl. 3:7195-7196, 2002. www.pnas.org/cgi/doi/10.1073/pnas.092078899.

Moss, S., Policy analysis from first principles, PNAS, 99:suppl. 3:7267-7274, 2002. www.pnas.org/cgi/doi/10.1073/pnas.092080699.

Carley, K.M., Computational organization science: A new frontier, PNAS, 99:suppl. 3:7257-7262, 2002. www.pnas.org/cgi/doi/10.1073/pnas.082080599.

Session 12: Artificial Chemistry

Dittrich, P., Jens Ziegler, Wolfgang Banzhaf. 2001. Artificial Chemistries – A Review, Artificial Life, Vol. 7 No. 3, pp 225-275.

Additional Reading:

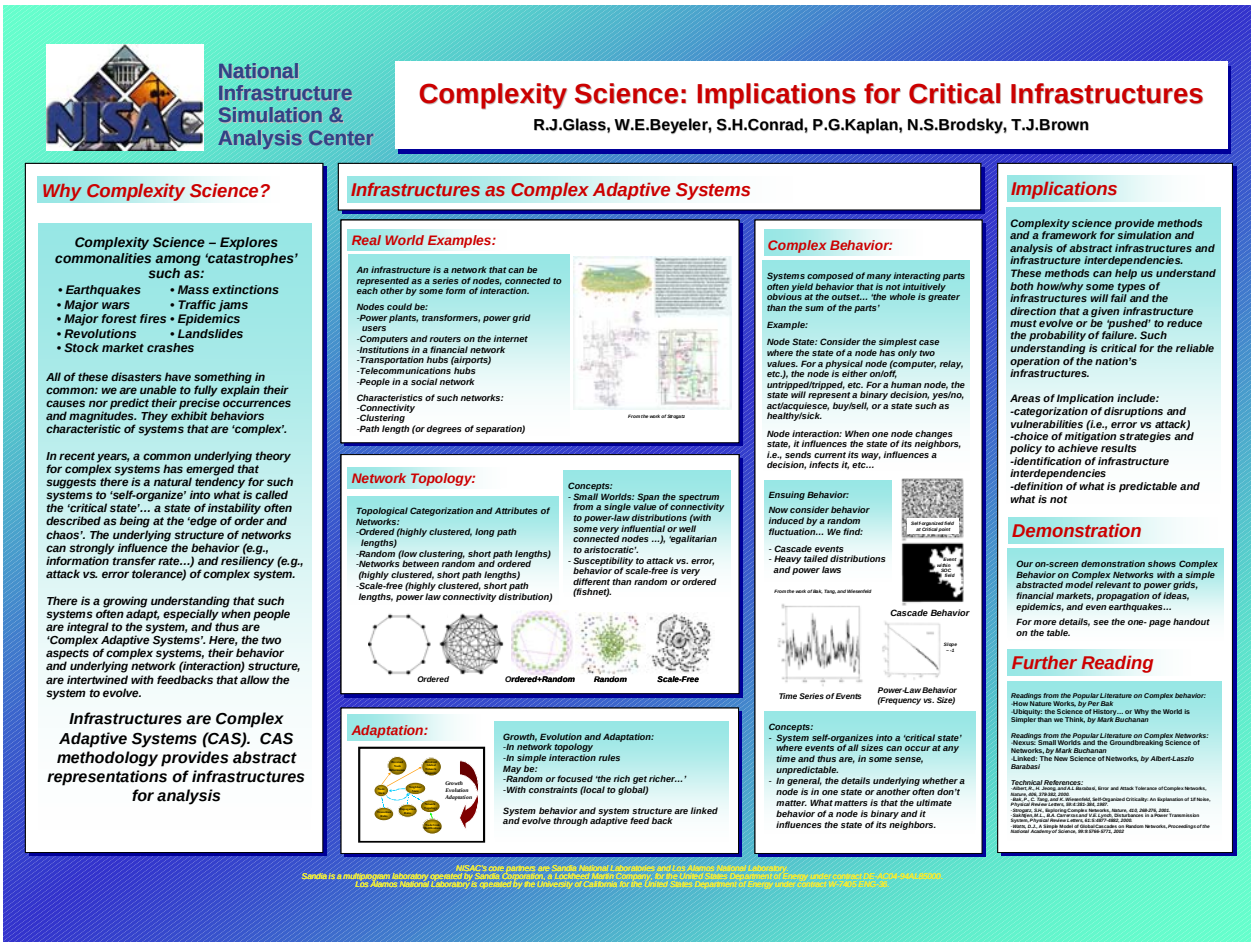
Kauffman, S. A. 1993. The origins of order: Self-organization and selection in evolution. New York, Oxford University Press.

Fontana, W., and L. W. Buss. 1996. The barrier of objects: From dynamical systems to bounded organization. In Boundaries and Barriers (pp 56-116). Redwood City, MA, Addison-Wesley.

Ziegler, J. and Wolfgang Banzhaf. 2001. Evolving Control Mechanisms for a Robot, Artificial Life, Vol. 7 No. 2, pp 171-190

Appendix B: Poster Presentation at NISAC's Capability Demonstration

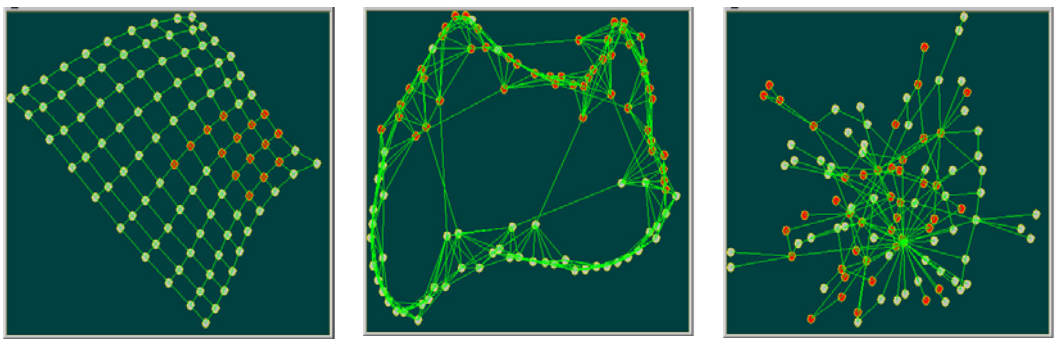
The Infrastructure Complexity R&D Group was asked present a poster as part of NISAC's Capability Demonstration in Portland, Oregon (March 26-27, 2003), and in Seattle, Washington (April 2-3, 2003). While this poster was not all inclusive of our work, it distills and communicates the essence of many of the concepts that we covered in this research activity. An interactive demonstration was part of the poster presentation and is outlined in Appendix C.



Appendix C: An Illustrative Example of Complex Behavior on Complex Networks: The "Generic Sand Pile"

Consider a Simple Abstracted Model with relevance to power grids, financial markets, propagation of ideas, epidemics, and even earthquakes… Patterned after the work of Bak, Tang, and Wiesenfeld [1987] and Sakhtjen, Carrerras and Lynch [2000].

Model: Each node has two states, passive and active. When in the passive state it does nothing. But when it is in the active state, it passes some of its 'Energy' (or information, stress…) to its neighbors, thereby influencing its neighbors' states as well. Each node has an Energy level, E, the value of which determines its state. When this variable exceeds a critical value, E_{crit}, the node changes its state from passive to active (on/off, tripped/untripped, buy/sell, express/acquiesce, slip/no-slip…). The behavioral rule for the active state is for a node to off-load a portion of its Energy to its neighbors. This is done randomly in units of one up to a prescribed portion E_{off}. This transfer of Energy (charge, information, etc) could cause the neighboring nodes to also exceed their critical values, and so on, thus creating a Cascade of activity. Individual nodes are connected to each other on a variety of grid topologies that span from regular grids, to random, and to those exhibiting 'small world' properties such as Watts-Strogatz or scale-free networks.



Red nodes are undergoing a Cascade event

Regular Square Lattice WS network Scale-free BA network

Now, let's begin with all nodes having Energy below E_{crit} but at different values given by a distribution with a mean value, $<E>$. We now perturb the system by selecting a random pair of connected nodes and push one increment of E from one to the other. We then determine if this shift causes the acquiring node to go above E_{crit} and thus change state. If it does, then we follow the off-loading of Energy to the neighboring nodes and any subsequent state changes that ensue.

Parametric variation: we can consider the influence of

- separation of average node state from critical value, $E_{crit}-<E>$
- Energy off-loaded by a node when it changes state, E_{off}
- mean connectivity of network, $<k>$
- network topology
- size of initial perturbation

Macro-system behavior: defined by the state change events in time, their sizes and statistics.

- ***cascade behavior*** on all general network topologies (regular, random, WS, scale free)
- ***fat tailed event distributions*** yielding power laws
- general dependence on the scaling parameter λ, where $\lambda=k^{\alpha}E_{off}/(E_{crit}-<E>)$
- occurrence of ***infinite cascades*** for λ above a critical value λ_{crit}
- details of when, where, and how big, ***are not predictable!***