Title: Predicting the Unpredictable in Complex Information Systems
Proposal Champions: Kevin Mills (772), James Filliben (776), Burt Rust (771), Chris Dabrowski (772), and Daniel Genin (772)

Summary: Modern society grows increasingly dependent on large information systems, such as the Internet and computational clouds, which consist of millions of components whose interactions lead to global dynamic patterns that cannot be predicted by analyzing the behavior of individual components. Such global patterns can include cascading failures, phase transitions and oscillations that drive systems from normal operating ranges to degenerate regimes that entail substantial cost. For example, one study [1] found that a 10-day Internet outage cost individual companies between $22M and $405M depending on the specific industry. When such outages affect major service providers [2], aggregate costs escalate quickly across affected companies to billions of dollars per hour.

System operators possess tools and procedures to mitigate degenerate behavioral regimes, but lack the ability to predict their onset. A recent *Nature* article [3] reports that "…certain generic symptoms may occur in a wide class of systems (e.g., physical, biological, geological and financial) as they approach a critical point." The generic symptoms relate to a systemic slowing down, which implies systems become increasingly slow in responding to small perturbations. The *Nature* article and others [4-6] describe a number of mathematical, statistical and computational techniques (e.g., eigenanalysis, autocorrelation and variance analysis, analysis of skewness and flickering) that might be applied to time series measurements to identify an approaching critical point. The articles also discuss selected patterns (e.g., scale-invariant power-law structures) that may appear spatially as systems approach criticality.

We propose to create algorithms combining mathematical, statistical and computational techniques that signal incipient changes in natural systems, and to evaluate those algorithms as a measurement basis to predict onset of phase transitions in large information systems. From previous research [7-9], we possess simulations of such information systems, which can be used to generate failure cascades, phase transitions and oscillations. We will evaluate the effectiveness of our algorithms at signaling phase changes in simulated spatiotemporal patterns. Subsequently, we will evaluate the most effective of our algorithms against data generated in real time with laboratory systems, such as the *Emulab* facility [10] installed within an ITL laboratory, and against data archived from real systems, such as the Internet [11-18] and other large distributed systems [19-20]. If we establish effective algorithms that can be deployed practically in real systems, then we will seek industrial collaborators to test the algorithms in operational systems or industrial test beds. Finally, if successful, we will advocate with designers and operators of large information systems to implement appropriate measurement methods to predict onset of degenerate regimes, which will reduce the frequency and scope of outages and save the U.S. economy billions of dollars per year. These savings may be amplified if applied to other critical infrastructures, such as the electric grid [21-22]. In addition, our research and findings will add to knowledge about detecting critical transitions in complex systems.

Context: Commerce, government, national utility grids and social interactions depend increasingly on large information systems, based on the global Internet. Disruptions of such systems, which appear likely to increase in spatial and temporal extent, incur significant economic costs for society. For example, Table 1 summarizes results from a 2003 telecommunications research study [23] to determine costs associated with network outages and

degradations for six specific companies in different industries. As shown, losses from either complete outages or periods of degraded performance cost companies as much as $100K/hour. Extrapolating to multiple companies, affected for extended periods by outages and degradations within large Internet service providers, implies staggering costs in aggregate. For example, a 2005 study [24] of outages across companies and industries estimated overall costs averaged about 3.6% of annual revenues. Such outages continue to occur. In 2007, a switchover to backup routes stimulated thousands of Cisco routers to rewrite routing tables in a short period, disconnecting millions of Internet users in eastern Japan for nearly seven hours [25].

Table 1. Summary of Infonetics Study of Network Downtime Costs in Six Companies [23]

Case Study	Revenue/Year	Downtime Cost	Cost/Hour	Outages	Degradations
Energy	$6.75 billion	$4.3 million	$1624	72%	28%
High Tech	$1.3 billion	$10.2 million	$4167	15%	85%
Health Care	$44 billion	$74.6 million	$96,632	33%	67%
Travel	$850 million	$2.4 million	$38,710	56%	44%
Finance	$4.0 billion	$10.6 million	$28,342	53%	47%

Beyond potential for large costs, network routing mishaps may also open the door for significant national security threats and malicious intrusions. In April 2010, for example, China Telecom advertised inaccurate traffic routes to the global Internet, leading to an 18-minute period during which as much as 15% of data crossing the Internet was forwarded through China [26]. A similar incident occurred in 2008, when routes advertised by a Pakistani Internet service provider caused all Internet traffic bound for YouTube to be sent to Pakistan [27].

In current practice, network operation centers [28] integrate numerous measurements, including utilization, status and alarms, from individual components. Usage data, e.g., link utilizations collected over five-minute intervals, reflect long-term trends that may signal need to reorganize network capacity. Ongoing component status is displayed on numerous large screens. Widespread outages and degradations are detected in real time via alarms arising independently from numerous network components. Alarms indicate significant system change has already occurred, and it can take hours [29] to discover underlying causes and begin corrective actions. We propose to investigate measurement techniques that signal approaching changes in global system dynamics before they occur with large magnitude over a wide scale. Advanced notice will enable system operators to begin problem isolation and resolution before such changes occur, shortening, or perhaps avoiding, outages. In this way, our work will advance measurement science to enhance economic security.

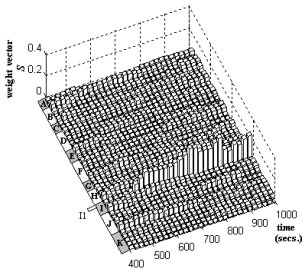
Technical plan: Our innovative idea is to apply techniques used to signal phase transitions in complex natural systems [3] to construct algorithms that predict incipient change in manmade, engineered systems. We will perform our research in the context of complex information systems, such as the Internet, computational clouds and wireless networks. This research is high risk because complex information systems: (1) encompass a large, multivariate, spatiotemporal scale, (2) exhibit highly correlated dynamics, (3) operate over multiple timescales, and (4) require real-time prediction of macroscopic behavior. This combination of traits suggests that individual techniques, such as change-point detection, eigenanalysis, autocorrelation analysis and analysis of flickering, could

prove insufficient. Successful approaches may require a combination of techniques to analyze correlation variation over extensive spatiotemporal, multivariate data sets that change on sub-second scale. If we succeed, our work would improve society's ability to measure and manage large systems. Whether we succeed or fail, our work would contribute a systematic framework to better understand techniques and methods to predict phase transitions in complex systems.

Past research has established that warning signals exist for phase transitions in simple systems, such as lasers [30] and neurons [31]. Evidence suggests that such warning signals exist in complex systems, such as glacial cycles [32-33], ecosystems [34-36], biological processes [37-38] and financial markets [39-40], though investigations continue. Researchers have also established the existence of phase transitions in traffic networks [41-43], electrical distribution networks [44-45], epidemics [46-47] and communication networks [48-50].

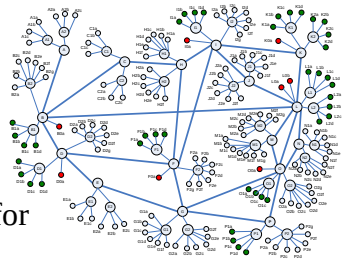
Our plan organizes into four phases: (1) Develop Theoretical Framework (18 months), (2) Conduct Theoretical Evaluation (12 months), (3) Conduct Practical Evaluation (30 months), and (4) Advocate and Disseminate (during and after the project). Major decision points arise following Phase 2 and 18 months into Phase 3. We address each phase in turn.

Phase 1–Develop Theoretical Framework. In this phase, we plan three activities: (1) to create algorithms by combining selected mathematical, statistical and computational techniques that can reveal early warning signals in complex natural systems, (2) to determine what measurement data are required and where and how to make the measurements in networks, and (3) to define criteria and procedures to evaluate our algorithms.



Activity 1.1 Techniques. Information networks exhibit correlated, multivariate data changing over large spatial extent and on multiple timescales. We will investigate combinations of techniques [3,51], such as change-point detection, eigenanalysis, autocorrelation and variance analysis and analysis of skewness and flickering, which can characterize temporal fluctuations and changes in variance and autocorrelation. We will also investigate techniques, such as random matrix theory [52], which can monitor changing spatiotemporal patterns. We will combine selected techniques into algorithms that can operate at the sub-second speeds with which network behaviors evolve. The main deliverables of this activity will be a set of algorithms that can be applied to measure global network dynamics and software codes that implement the required computations.

Activity 1.2 Simulation. We will identify network behaviors that exhibit phase changes and determine the measurements required to monitor them. We will assess where and how such measurements can be simulated within network models [e.g., 7-9]. We will determine specific network models to use when evaluating change-detection algorithms. We will parameterize the chosen models to generate classes of phase transitions of interest, and demonstrate that ability. We will instrument simulations to collect measurements required to inform the algorithms created in Activity 1.1. The main deliverables of this activity will be a set of network simulation models capable of generating phase transitions of interest and of collecting data required for algorithmic analysis.

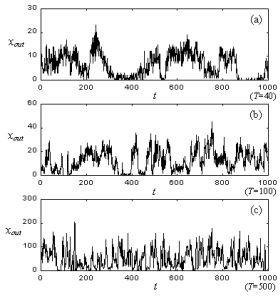


Activity 1.3 Criteria. We will define procedures to evaluate the algorithms created during Activity 1.1, when attempting to achieve early detection of phase changes simulated in Activity 1.2. We will also define criteria and metrics to evaluate the degree of success, including analysis

of false positives and negatives, achieved by our algorithms, when used to detect onset of network phase changes. The main deliverables of this activity will be criteria, metrics and procedures that we will use for theoretical evaluation during Phase 2.

Phase 2–Conduct Theoretical Evaluation. In this phase, we will execute the evaluation plan developed during Activity 1.3 in the context of the simulation models and phase transitions created during Activity 1.2. We will simulate the chosen phase changes with our network models, and collect data necessary to evaluate the algorithms created in Activity 1.1. We will characterize the ability of our algorithms to detect macroscopic changes in network dynamics. We will assess the extent of success along various dimensions, such as statistical reliability, detection latency, speed of processing, applicability to real networks, and other criteria defined during Activity 1.3. The main deliverable from this activity is a candidate set of algorithms to forward to the next phase of evaluation. Should none of our algorithms pass the minimal success criteria identified in Activity 1.3, the project should be reevaluated, possibly returning to Phase 1 to consider alternate algorithms, or being canceled, if no promising algorithms exist.

Phase 3–Conduct Practical Evaluation. We enter Phase 3 with a set of promising algorithms to detect incipient phase changes in information networks. The algorithms passed a rigorous examination in simulated settings, but at this point must be subjected to more realistic testing. Establishing industry collaborations will be crucial to achieve sufficient realism. As we enter Phase 3, we will issue a *Federal Register* notice seeking industry collaborators to establish a research and development agreement (CRADA) and consortium. The main objective of Phase 3 is to establish whether or not the promising algorithms are practical and effective in the context of data from real networks. Achieving this objective requires pursuing three activities: (1) evaluate the algorithms on measurement traces adapted from archived data collected by other research groups, (2) evaluate the algorithms in real time on measurement data collected during empirical experiments conducted with available cloud-computing facilities, and (3) establish a consortium to test the algorithms in commercial systems or industrial laboratories. If no algorithms appear practical and effective after evaluation against trace-based and laboratory data, then the project should be reevaluated for possible return to Phase 1, or termination.

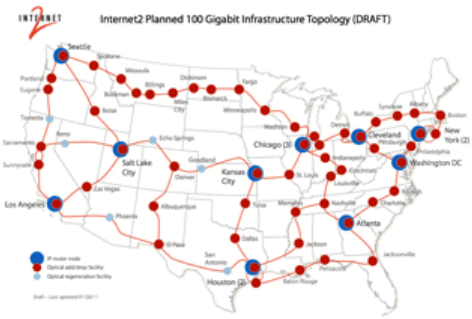


Activity 3.1 Trace-Based Evaluation. We will evaluate algorithms surviving from Phase 2 against behavioral traces adapted from archived data collected by other research groups [e.g., 11-20]. Since the collected data are unlikely to exhibit the phase transitions successfully signaled in Phase 2, we plan to generate phase changes by adapting the available traces. While such artificial generation has not been attempted previously, insights gained from Phase 2 should enable us to generate detectable phase changes and the archived data should provide realistic details. The main reason for taking this step is to achieve a reasonably economical evaluation of the surviving detection algorithms against realistic data, prior to evaluating the algorithms in real time using more expensive procedures, as foreseen in Activity 3.2. The main deliverable from this activity is a set of detection algorithms that are successful enough for laboratory evaluation in real-time usage.

Activity 3.2 Laboratory Evaluation. We will configure empirical network topologies to generate phase changes successfully detected in Activity 3.1. We will instrument the topologies to collect necessary measurement and evaluation data. We will transform our detection algorithms into software codes required to execute in real time in a laboratory network. We will evaluate the degree of success achieved by each surviving algorithm. We will scale our empirical

topologies to large size, using Emulab-based facilities, which exist in universities [53-54] throughout the world, or commercial cloud-computing facilities, such as the Amazon Elastic Compute Cloud [55], where access to virtual resources is nearly unlimited, provided we have sufficient funds. To minimize the cost of using external facilities, we will first prototype experiments on a local Emulab [10] installed within an ITL laboratory. The ITL Emulab has 60 physical nodes that can emulate about 1000 virtual nodes. With IE funds from this project, we intend to double the capacity of the ITL Emulab. Prototyping our experiments locally will ensure that they can execute successfully prior to committing resources to larger-scale experiments on external facilities. The main deliverables of this activity include a set of detection algorithms that are viable candidates for real-world use, software codes that can execute the algorithms in real time in operational networks, and instrumentation software to collect the data used by the algorithms.

Activity 3.3 Real-World Evaluation. Upon reaching this point, the project has created a set of detection algorithms (and codes) and monitoring software that have successfully signaled the onset of selected phase changes in real time, so it stands to reason that commercial concerns operating large networks should be interested in collaborating to further test our algorithms, either in an operational network or in an industrial laboratory. Some Internet service providers are becoming more open about their own measurement and management systems. Comcast, for example, issued a request for comments (RFC 6057) [56] that explains publically their approach to detecting and managing the onset of congestion caused by individual users in an access network. Such openness suggests that Internet service providers, such as Comcast, might be receptive to collaborate on techniques that detect incipient onset of spatiotemporal congestion within a network at large. Upon entering Phase 3, we will invite all interested collaborators to enter a CRADA consortium. The main practical hurdle to overcome is constructing a deployment scenario that has no possibility of disrupting a company's system. We expect to evade such problems by constructing, during Activity 3.2, passive approaches to measure the required data. If data collection entails modifying operational components, then we are less likely to convince companies to test our techniques in an operational network, which means we would have to collaborate with companies that possess laboratories for experimenting with novel approaches. The main deliverable from this activity will be collaborative evaluation of selected detection algorithms in the context of one or more real systems, or industrial laboratories.



Phase 4–Advocate and Disseminate. Phase 4 can take several directions, depending on our degree of success. If fully successful, we will advocate at industry groups, such as the North American Network Operators Group (NANOG), for industry to adopt our algorithms and to join our CRADA consortium to investigate and evaluate additional algorithms. Further, we will present our results to industry groups with related complex systems, such as utility grids and transportation networks. If only partially successful, then our insights published in papers during the course of the research will contribute a systematic framework to better understand techniques and methods to predict phase changes in complex information systems. If we fail, then reports from our research will provide a foundation for others to move forward better informed. The

main deliverables from this phase include the papers, presentations and reports produced during our research, which will close with a final report documenting what we did, the degree to which we succeeded, and suggested directions for future research.

Potential impacts: Studies [1,23-24] show that network outages and degradations can cost individual companies as much as $100K/hour, which can expand to $1B/hour, when the affected network serves 10,000 companies. Current practice alerts network operators after outages begin, which typically leads to hours taken to diagnose and correct the problem. A single outage, then, can cost billions of dollars, and when aggregated across all system outages and degradations, the societal cost becomes staggering. If we succeed, then network operators can be alerted to many incipient problems before they occur, allowing early diagnosis and correction, and leading to reduced costs. In some cases, effective detection may enable operators to mitigate degradations and avoid associated costs. Further, our work will contribute to emerging scientific understanding of the nature of phase transitions in complex systems. Through this project, NIST can become the leading organization researching phase-change detection in complex information systems. Our findings may also inform research on phase changes and cascading failures in other complex engineered systems, such as utility grids.

Qualifications of research team: Expertise of the research team spans mathematics (Rust and Genin), statistics (Filliben), and computer science (Mills and Dabrowski).
(1) **Kevin Mills** has created simulation models and measurement codes for large communication networks, computational grids and clouds. He has designed statistical experiments to reveal the macroscopic behavior of congestion control algorithms and resource allocation schemes. He has developed performance measurement code deployed into operational networks.
(2) **Jim Filliben** possesses unique skills in experiment design, exploratory data analysis and time series analysis that he has applied to a diverse set of problems spanning the physical sciences. He has recently developed and applied methods to analyze global behavior in complex networks.
(3) **Burt Rust** is a world-class expert in developing mathematical models to characterize time series data, and has applied that expertise to model data ranging from climate change measurements to traffic measurements from the Internet.
(4) **Chris Dabrowski** has created Markov chain models and simulations representing large distributed systems, including infrastructure clouds, computational grids and communication networks. He recently developed methods to predict causes and expected patterns of performance degradation in large distributed systems.
(5) **Daniel Genin** joined NIST as an NRC post-doc in 2006, after which he developed analytical models improving the accuracy of fluid-flow approximations intended to predict the behavior of congestion control algorithms for the Internet. In 2010, he joined the regular staff and began analyzing resource allocation algorithms for computational clouds.

Resources required: Removed from this version.

References:

[1] S. Dynes, E. Andrijcic, and M. E. Johnson, "Costs to the U.S. Economy of Information Infrastructure Failures: Estimates from Field Studies and Economic Data", *Proceedings of the 5th Workshop on the Economics of Information Security*, Cambridge University.

[2] R. Charette, "Comcast Suffers Major Internet Outage", *IEEE Spectrum*, Nov. 29, 2010.

[3] M. Scheffer et al., "Early warning signals for critical transitions", *Nature*, v. 461, no. 3, Sep. 3, 2009, pp. 53-59.

[4] J. Lim and B. Epureanu, "Forecasting a class of bifurcations: Theory and experiment", *PHYSICAL REVIEW E* 83, 016203, 2011.

[5] C. Kuehn, "A mathematical framework for critical transitions: normal forms, variance and applications" Cornell University Library, arXiv eprint, 1101.2908. arXiv:1101.2908v1 [math.DS]

[6] R. Washington-Allen, D. Briske, H. Shugart, and L. Salo, "Introduction to special feature on catastrophic thresholds, perspectives, definitions, and application", *Ecology and Society*, 15(3): 38, 2008.

[7] K. Mills, E. Schwartz and J. Yuan, "How to model a TCP/IP network using only 20 parameters", *Proceedings of the 41st Winter Simulation Conference*, 2010, IEEE, 849-860.

[8] K. Mills, J. Filliben and C. Dabrowski, "Sensitivity Analysis of Koala: an Infrastructure Cloud Simulator", submitted to the *4th International Conference on Cloud Computing*, IEEE, 2011.

[9] K. Mills and C. Dabrowski, "Can Economics-based Resource Allocation Prove Effective in a Computation Marketplace?", *Journal of Grid Computing*, 6/3, September 2008, 291-311.

[10] B. White, J. Lepreau, L. Stoller, R. Ricci, S. Guruprasad, M. Newbold, M. Hibler, C. Barb, and A., Joglekar, , "An Integrated Experimental Environment for Distributed Systems and Networks," *Proceedings of the 5th Symposium on Operating Systems Design and Implementation*, 2002, 255-270.

[11] The Internet 2 Observatory Data Collections, http://www.internet2.edu/observatory/archive/data-collections.html

[12] Protected Repository for the Defense of Infrastructure against Cyber Threats, https://www.predict.org/Default.aspx?tabid=40

[13] The Cooperative Association for Internet Data Analysis (CAIDA) Data Sets, http://www.caida.org/data/overview/

[14] RIPE Network Coordination Centre Data Sets, http://labs.ripe.net/datarepository/data-sets

[15] K. Heidemann and C. Papdopoulos, "Uses and Challenges for Network Datasets", *Proceedings of Conference for Homeland Security*, 2009, 73-82.

[16] University of Illinois Urbana-Champaign Repository of Availability Traces http://www.cs.uiuc.edu/homes/pbg/availability/

[17] Network Economics Group Data Sets, http://netecon_group.tmit.bme.hu/source-codes

[18] University of California Riverside Archive, http://networks.cs.ucr.edu/ucrchive/measurement.htm

[19] Google M-Lab Data Sets, http://www.measurementlab.net/data

[20] Technische Universiteit Delft, Peer-to-Peer Trace Archive, http://p2pta.ewi.tudelft.nl/pmwiki/?n=Main.Home

[21] D. Newman, B. Carreras, V. Lynch, and I. Dobson, "Exploring complex systems aspects of blackout risk and mitigation", *IEEE Transactions on Reliability*, in press, 2011.
[22] P. Hines, E. Cotilla-Sanchez, and S. Blumsack, "Topological Models and Critical Slowing Down: Two Approaches to Power System Blackout Risk Analysis", *Proceedings of the 44th Hawaii International Conference on System Sciences*, 2011.
[23] J. Wilson, "The cost of network downtime 2003", Infonetics Research.
[24] J. Wilson, "The cost of network downtime 2005", Infonetics Research.
[25] J. Duffy, "Cisco routers caused major outage in Japan", Network World, May 16, 2007.
[26] J. Cowie, "China's 18-Minute Mystery" Renesys Blog, November 18, 2010. http://www.renesys.com/blog/2010/11/chinas-18-minute-mystery.shtml
[27] M. Brown, "Pakistan hijacks YouTube", Renesys Blog, February, 24, 2008.
[28] ASIA/PAC AMHS/ATN Network Management Operational Procedure Guidelines, International Aviation Organization, September, 2009.
[29] D. Peng, "About the April 13th Service Outage", Ooma's Blog, April 2009.
[30] J. Tredice et al., "Critical slowing down at a bifurcation", *American Journal of Physics* 72, 799-809, 2004.
[31] G. Matsumoto and T. Kunisawa, "Critical slowing-down near the transition region from the resting to time-ordered states in squid giant axons", *Journal of the Physical Society of Japan*, 44, 1047-1048, 1978.
[32] J. Petit et al. "Climate and atmospheric history of the past 420,000 years from the Vostok ice core, Antarctica", *Nature*, 399, 429-436, 1999.
[33] D. Luthi et al., High-resolution carbon dioxide concentration record 650,000-800,000 years before present", *Nature* 453, 379-382, 2008.
[34] M. Scheffer and S. Carpenter, "Catastrophic regime shifts in ecosystems: linking theory to observation", *Trends in Ecology and Evolution*, 18, 648-656, 2003.
[35] M. Rietkerk et al., "Self-organized patchiness and catastrophic shifts in ecosystems", *Science*, 305, 1926-1929, 2004.
[36] M. Scheffer and E. van Nes, "Shallow lakes theory revisited: various alternative regimes driven by climate, nutrients, depth and lake size", *Hydrobiologia*, 584, 455-466, 2007.
[37] J. Venegas et al., "Self-organized patchiness in asthma as a prelude to catastrophic shifts", *Nature*, 434, 777-782, 2005.
[38] F. Mormann et al. "Seizure prediction: the long and winding road", *Brain*, 130, 314-333, 2007.
[39] D. Bates, "The crash of 87 – was it expected? The evidence from options markets", *Journal of Finance*, 46, 1009-1044, 1991.
[40] R. Whaley, "Derivatives on market volatility: hedging tools long overdue", *Journal of Derivatives*, 1, 71-84, 1993.
[41] J. Cuesta et al. "Phase transitions in two-dimensional traffic-flow models", *Physical Review E*, 48, 4175-4178, 1993.
[42] D. Helbing et al., "Micro- and macro-simulation of freeway traffic", *Mathematical and Computer Modeling*, 35, 517-547, 2002.
[43] Y. Yokoya, "Dynamics of traffic flow with real-time traffic information", *Physical Review E*, 69, 11 pages, 2004.
[44] B. Carreras et al., "Critical points and transitions in an electric power transmission model for cascading failure blackouts", *Chaos*, 12:4, 10 pages, 2002.

[45] H. Liao, J. Apt and S. Talukdar, "Phase Transitions in the Probability of Cascading Failures", Carnegie Mellon Electricity Industry Center, 2006.
[46] Y. Moreno, R. Pastor-Satorras and A. Vespignani, "Epidemic outbreaks in complex heterogeneous networks", *European Physical Journal B*, 26:4, 521-529, 2004.
[47] T. Zhou, Z-Q. Fu, B-H Wang, "Epidemic dynamics on complex networks", *Progress in Natural Science*, 16:5, 452-457, 2006.
[48] K. Fukuda, H. Takayasu, and M. Takayasu, "Origin of critical behavior in Ethernet traffic", Physica A, 287, 289-301, 2000.
[49] B. Krishnamachan, S. Wicker and R. Bejar, "Phase transition phenomena in wireless ad hoc networks", *Proceedings of the Global Telecommunications Conference*, 5, 2921-2925, 2001.
[50] E. Coffman et al. "Network Resilience: Exploring Cascading Failures with BGP", *Proceedings of the 40th Allerton Conference on Communication Control and Computing*, 40, 97-106, 2002.
[51] H. Wang, D. Zhang and K. Shin, "Change-point monitoring for the detection of DoS attacks", *IEEE Transactions on Dependable and Secure Computing*, 1:4, 193-208, 2004.
[52] J. Yuan and K. Mills, "Monitoring the Macroscopic Effect of DDoS Flooding Attacks", *IEEE Transactions on Dependable and Secure Computing*, 2:4, 324-335, 2005.
[53] Emulab total network testbed http://www.emulab.net/
[54] Other Emulabs https://users.emulab.net/trac/emulab/wiki/OtherEmulabs
[55] Amazon Elastic Compute Cloud (EC2) http://aws.amazon.com/ec2/
[56] C. Bastian et al. "Comcast's Protocol-Agnostic Congestion Management System", RFC 6057, December 2010.

Title: Predicting the Unpredictable in Complex Information Systems - Response to Questions
Proposal Champions: Kevin Mills (772), James Filliben (776), Burt Rust (771), Chris Dabrowski (772), and Daniel Genin (772)

Question 1: *The authors mention critical slowing down as a tool for identifying system behavior near a critical point. Many phase transitions in natural systems are first order, and do not exhibit this precursor. How confident are the authors that the techniques they wish to use will cover the relevant transitions?*

We are interested in detecting shifts in the dynamic behavior of information systems, such as computer networks, where dynamic behavior can be represented as spatiotemporal patterns of network packets or router states. Some evidence for the existence of such shifts has already been suggested. For example, Takayasu and colleagues [1] observed data on a link connecting Keio University to public Japanese Internet backbones (known as WIDE) and discovered existence of a phase transition between sparse and congested traffic patterns. These researchers report that the macroscopic behavior of this phase transition exhibits characteristics similar to second-order phase transitions in physical systems. Unlike ordinary physical systems, the observed phase transition in Internet traffic appears to be dynamical, with the control parameter fluctuating slowly around a critical point. Other researchers [2-3] have identified both first- and second-order phase transitions in network models. Echenique and colleagues [2] suggest that, depending on the underlying routing protocol, networks can exhibit either first- or second-order transitions between congested and sparse traffic regimes. These researchers find that second-order transitions occur when routing is based on shortest path, as is the common case in today's networks. Even in the first-order case examined by these researchers, traffic flow in the network reorganizes itself into impermeable regions over a long time prior to a sudden jump in network-wide congestion, so there are likely to be signs, such as changing network traffic patterns, that can be detected as precursors to a first-order-like transition from free-flowing to congested traffic. Yuan and Mills [4], for example, applied random-matrix theory to construct a measurable signal of shifting traffic patterns arising from various distributed denial of service attacks, where even the stealthiest attacks could be detected prior to sudden onset (i.e., first-order transition) of access link congestion.

Coffman and colleagues [5], when exploring phase transitions in cascading failures among border-gateway protocol (BGP) routers, do not adopt the distinction of first- and second-order phase transitions, and instead define (ala Erdos and Renyi) a phase transition as an abrupt change in a system-wide property. These researchers found that a collection of BGP routers can fail to recover from router crashes if the number of failed routers in the collection exceeds a critical threshold. Even in this case, one suspects that by monitoring the time varying distribution of crashed routers, a macroscopic signal can be constructed to predict the possible onset of network collapse. For example, Liao and colleagues [6] study cascading failures in electric power networks and suggest that online monitoring may possibly enable detection of approaching phase transition in the probability of the onset of cascading failures. Similarly, Carreras and colleagues [7] examined a 15-year time series of system blackouts in the North American power grid and discovered that the probability distribution of blackout sizes exhibits a power law, suggesting that the electrical grid has historically been operated near a critical point. These researchers identify two different phase transitions: one second-order (based on increased power shedding

with increasing demand) and one first-order (based on limitations in individual power lines). The changing rate of power shedding could certainly be the subject of a monitoring regime aimed at predicting a transition in the probability of blackout cascades.

In summary, the techniques we plan to investigate can be applied to measure spatiotemporal shifts in the pattern of network traffic or element states. Whether or not the actual phase transitions we aim to detect appear as first- or second-order transitions, our hypothesis is that, in many useful cases, such transitions will be preceded by precursor shifts in macroscopic patterns of system-wide dynamics or state. Above, we identified some investigations with respect to network traffic, router state and electrical power-shedding dynamics that provide evidence supporting our hypothesis. Of course, the research we propose is high risk, and so our hypothesis may turn out to be wrong, and even if our hypothesis is correct, we may be unable to construct reasonable pattern-detection regimes that can be applied practically with the necessary granularity and timeliness. Our confidence of success stems in part from a past success [4] in detecting stealthy distributed denial of service attacks by constructing a signal based on random-matrix theory. We believe that similar approaches can be developed to expand the range of detectors available to forecast other anomalous behaviors.

Question 2: *The method discussed in the proposal, based on the study of natural systems, is well suited to intrinsic changes in large-scale behavior arising from interactions within the system, but many historically noteworthy complex-system failures have arisen from extrinsic events, such as component failure or operator error. Can the proposed analysis scheme identify vulnerabilities to any generic class of extrinsic events, such as unanticipated edge-cutting or node failure?*

As we explained above, our proposed research is motivated by a desire to detect macroscopic spatiotemporal shifts in system dynamics or element state that suggest an incipient change in the probability of moving to an anomalous regime. The cause of such shifts may be extrinsic events, such as component or link failures [5-7], or may occur as part of intrinsic system dynamics, such as variations in network traffic, arrival of flash crowds and stealthy distributed denial of service attacks [1-4]. Our hypothesis is that most engineered systems, such as computer networks or electric power grids, operate at a highly efficient dynamic point, where the system is near a critical transition boundary between optimal and highly suboptimal operating regimes. At such points, slight perturbations in (intrinsic) system load and/or (extrinsic) available resources can tip the system into one of the available suboptimal regimes. We suspect that engineered systems exhibit a characteristic slowing down [8, 26] as such critical points are approached. If we are able to detect and signal situations where a system approaches a critical point, then operators should be able to take steps to restrain the system from reaching the critical point, thus lowering the probability of transitioning to a suboptimal regime. If our hypothesis is correct, and if we can develop practical measurement and signaling techniques, then we should be able to provide alerts that a system is approaching a critical point. Further, we might also be able to estimate the distance between current system operating state and such a critical point. If we can accomplish this, then we might also be able to provide a time-varying map showing changes in distance between a systems' operational state and its critical state. If we can achieve such an outcome, then system operators might be able to experiment with configuration changes and observe their effect on the distance between system state and a critical state. Such an outcome, though

certainly more than we expect to achieve, could open up complex information systems to the possibility of dynamic steering among tradeoffs in performance and robustness.

On the other hand, we do not expect to develop methods that can necessarily detect the approach of every possible anomalous behavior caused by extrinsic events. For example, a human operator could develop and deploy an incorrectly executed configuration file that causes network traffic to be instantaneously misrouted as soon as the configuration is activated. The onset of such an anomalous regime might well occur so quickly that the type of pattern-detection methods we envision would have insufficient time to detect spatiotemporal shifts. Such an erroneous configuration would effectively set a network immediately into an anomalous regime. We envision developing methods that detect shifting spatiotemporal patterns, and not discontinuous changes that occur immediately as a result of such events as erroneously generated network configurations.

In summary, we seek techniques to detect changes in macroscopic patterns of system behavior over time and space, regardless of the source, which may include causes both intrinsic (e.g., changes in demand patterns) and extrinsic (e.g., changes in resource availability due to component failures). In our experience, the operating physics of large information systems stems from an underlying relationship between demand and available resources. Changes in either demand or resource availability (or both) will influence this relationship and contribute to changing spatiotemporal patterns of system behavior. We plan to focus on detection techniques and analysis schemes that apply regardless of the sources of system disturbance, as long as such disturbances unfold in space over a period of time. Most system disturbances do unfold over space and time, but some (such as erroneously generated system reconfigurations) may occur immediately, and so be unpredictable using the methods we envision.

Question 3: *Computer system operators have an obvious motivation to work to maintain the reliability of their networks. How does the proposed effort compare to similar efforts already underway in industry?*

Our review found that industry has its main focus on network engineering, network security and traffic monitoring for capacity planning. There appear to be no current industry initiatives to research detection of phase transitions for application in networks and other distributed information systems. As explained in our proposal, system operators generally operate today on two time scales: (1) responding to real-time alarms generated by system management monitors and (2) planning capacity expansion based on monitoring long-term trends in usage. On the real-time scale, in the case of many alarms occurring closely in time, we are arguing that system dynamics have already crossed (or will very soon) a critical threshold, driving the system into a suboptimal regime from which it must be recovered. Our proposal cites several such cases. Even on the long-term scale, network operators today appear to be trading off capacity expansion against heuristic bandwidth management techniques at the network edge, leading to statistical multiplexing based on rules of thumb, such as “one can support 1000 100 Megabit per second users with a 1 Gigabit per second link” [9]. Adopting such heuristics seems likely to drive a network closer to a critical threshold. The research we are proposing addresses an intermediate range of time scales, where system dynamics evolve prior to crossing over into a suboptimal regime.

Several of our colleagues attend quarterly meetings of network operators, such as NANOG, the North American Network Operators Group [10]. Our colleagues report that the topic of phase transitions in networks is not discussed at those meetings. A perusal of the presentations at NANOG reveals an emphasis on engineering networks, securing networks, monitoring traffic trends and analyzing network incidents of various kinds. Our colleagues who attend the Internet Engineering Task Force (IETF) [11] report a similar lack of discussion regarding phase transitions in networks. At the IETF there is some work (ConEx) aimed at introducing management mechanisms that can allow network operators to expose and restrict users who are the source of network congestion. This work addresses congestion as a somewhat localized issue, without considering its network-wide implications.

About the closest set of industry research that we could find related (somewhat) to our proposal centers around intrusion-detection systems, as funded by DARPA circa 2000 [12]. The challenge in intrusion-detection research is to devise techniques (some statistical and some signature-based) to detect the onset of cyber attacks. In most cases, the cyber attacks to be detected are aimed at particular host computers or links on a network edge, such as a campus or corporate site, rather than within the network as a whole. Some detection techniques look at behavior on computers and some examine network traffic entering and leaving computers. In the DARPA program, each of eighteen detection schemes (some from industry and some from academe) was exposed to 200 instances of 58 types of attacks embedded within a live background of traffic related to hundreds of users operating on thousands of computers. Some detection methods generated a high level of false alarms, while others generated under 10 false alarms per day. The performance of the detectors that had low false-alarm rates was generally poor otherwise, detecting around 50 % of attacks that were brute force in nature and under 20 % of more stealthy attacks. In fact, 10 attack types were not detected at all, by any method. Our proposal already recognizes the importance of minimizing false alarms, which we identify as one criterion in the evaluation of proposed techniques. Further, unlike intrusion-detection systems, circa 2000, we already have identified one technique to detect stealthy distributed denial of service attacks [4]. Finally, unlike intrusion-detection systems that focus on network access links, we intend to develop detection techniques that operate on network-wide behavior, where cyber attacks of interest include epidemic-style [27] spread of so-called Internet worms [28] that infect computers throughout a network.

In summary, industry has its main focus on network engineering, network security and traffic monitoring for capacity planning. There appear to be no current industry initiatives to research detection of phase transitions in networks and other distributed information systems. We hope that success on our proposed project will succeed in putting measurement science for detecting phase transitions on industry's research agenda.

Question 4: *Will the authors be able to take their algorithms and retroactively predict historic data that might exhibit these phase transitions, as a means of validating their models?*

In principle, the answer to this question is yes. On the other hand, to practically complete this type of validation, we need access to data the exhibits suitable phase transitions. Such data appears to be in short supply, at least from public sources. For this reason, we include in our research plan, a technical task (Activity 3.1) that evaluates our approaches against phase

transitions generated artificially from publically available traces of data [e.g., 13-22] collected under normal system operating regimes. Taking this step provides one means to validate our techniques, models and analyses.

There appear to be a limited number of measurement datasets, used in some studies of phase transitions, which we could exploit to help validate our approaches. For example Takayasu and colleagues [1] based their analysis on nine, four-hour, datasets collected from an access link connecting Keio University with WIDE. The data was converted into counts of bytes in packets in each 100 ms interval. We have inquired (no response yet) with Takayasu and his colleagues about the availability of that data. Similarly, Carreras and colleagues [7] have analyzed 15 years of data on cascading failure blackouts in the North American power grid. Next month, we will have an opportunity to meet with Ian Dobson (one of Carreras' coauthors) at a complex systems symposium and to inquire about the availability of that dataset. We will continue to search for publically accessible data exhibiting phase transitions in large distributed systems. We envision contacting researchers at organizations investigating complex systems, and also probing attendees at related conferences. If we can gain access to historical datasets containing phase transitions, then we could apply our proposed techniques to such data.

Beyond measurement datasets, there appear to be a wide collection of possible simulation data sources generated by various network models, such as those we identified here [2-5] and elsewhere [23]. Using such data would be analogous to validating the ability of our techniques, models and analyses to detect phase transitions in simulations, which is already included (Phase 2) in our work plan.

In summary, we found a dearth of publically available data exhibiting phase transitions. We are attempting to contact researchers known to have had access to such data. The general lack of historical data exhibiting phase transitions led us to propose validating our techniques, models and analyses against simulated data and against data adapted artificially from publically available traces of normal system operation. We believe the validation methods we propose will prove sufficient.

Question 5: *A study of this type could reveal denial-of-service vulnerabilities in active commercial systems. How confident are the authors that commercial system operators will be willing to share sufficient operational data (both historical data sets and sensor deployment opportunities) to test the proposed detection algorithms on live, real systems?*

In general, we have relatively low confidence that commercial system operators will be willing to share sufficient (historical) operational datasets with us. We have some limited experience in attempting to obtain such data from Internet Service Providers (ISPs). In fact, we have been successful in only one such case, where we had close personal contact with an executive at an ISP. Even in this case, the data was shared with strict requirements for encryption and with requirements that any use of the data in publications would be anonymized and cleared by the executive that authorized our access to the data. This experience leads us to conclude that, in general, we must have low confidence in obtaining such data.

Our confidence is somewhat higher with regard to sensor deployment opportunities, especially if such sensors can be limited to passive data collection. We assume that such opportunities can be undertaken only in the context of a collaborative research and development agreement (CRADA) that spells out the rights and responsibilities of all involved parties. We intend (see Activity 3.3) to attempt to establish such CRADAs. To have any chance to entice CRADA participants from industry, we must be able to demonstrate that our software is safe and our algorithms successful in a real-time setting similar to an actual network. We aim to fulfill this goal (see Activity 3.2) by evaluating our techniques in a realistically scaled laboratory network, first at NIST and then in a computational Cloud.

One year ago, we were quite pessimistic regarding engaging ISPs in live experiments. In the intervening time, as we discuss in our proposal, one ISP (COMCAST) has published the technical specifications of their approach to detecting and managing congestion. This fact increases our hopes that ISPs might be more willing to engage in experiments of the nature we have in mind. On the other hand, NIST is a government agency, which may be the type of organization that ISPs are least willing to work with. This stems in part from the fact that several other government agencies (such as DHS, FCC and NTIA) take steps to regulate various activities of network providers. Even NIST sets government-wide standards with respect to computer security.

In summary, the issue of interacting closely with ISPs is indeed a tricky one. Clearly, engaging ISPs will represent an uphill fight for us. We do have colleagues with long-standing involvement in NANOG and the IETF. These colleagues are well-respected and should at least allow us to gain a hearing with potential collaborators. In order to have a chance to engage ISPs, we will need to achieve all of the objectives we outlined leading up to Activity 3.3. Even in that case, we may still have a difficult challenge to engage ISPs to the extent we envision. Perhaps we will be better able to succeed in interesting ISPs to take up (and/or extend) our methods within their own research programs. Recall from our answer to **Question 3** that industry does not currently appear to be addressing the issue of phase transitions in networks – so if we can get as far as to inspire and inform internal industry efforts to research this topic and related detection schemes, then we will still have succeeded in engaging industry, if perhaps less directly than we hoped. In the end, having industry set its own research agenda related to predicting phase transitions would be a successful outcome.

Question 6: *Information systems tend to be interconnected in a variety of ways. Do the authors have a sense of where a reasonable boundary of an information system might lie? Is it sufficient or useful to instrument a single ISP's network, for instance, or is it necessary to instrument the worldwide Internet in order to obtain the expected predictive capability?*

This question can best be considered along two dimensions: (1) where must instrumentation be placed and (2) what measurements must be made. Of course, requiring that the entire worldwide Internet be instrumented and measured would be impractical, both from the standpoint of the volume of data to be collected and the cost of processing such data. While concrete answers on instrument placement and measurement collection will depend on the findings of our research, we can draw some insight from a previous investigation [4] into macroscopic detection of distributed denial of service attacks within a simulated four-tier ISP network with 11 backbone

routers, 40 point-of-presence routers, 110 access routers and 22,000 host computers. When modeling this network we envisioned deploying measurement instruments at any of the 40 point-of-present (second-tier) routers, and counting packets flowing outbound from each measurement point to all access (third-tier) routers. Our experiments found that we needed to make measurements at 10 % (i.e., four) of the 40 possible measurement points in order to observe sufficient traffic to detect shifting spatiotemporal patterns. Our ability to discern shifting traffic patterns diminished when we dropped the number of measurement points to two. Since we were interested in measuring traffic inbound to access routers, we needed to count packets in 110 bins (one per access router), so each measurement point collected 110 numbers every measurement interval, and periodically forwarded the collected data to a central point for processing. At the central point, we applied eignenanalysis to square (440x440) matrices to compute a weight vector, which acted as a signal of traffic intensity inbound to the access routers. In general, the larger the square matrix that must be analyzed, the more processing time is required. We determined that tradeoffs exist among the number of monitoring points, the number of measurements and the monitoring cycle time. We were able to show that we could successfully monitor shifting traffic at 10 monitoring points outbound for 110 access routers within a 90 s cycle time. Based on our analysis, we were encouraged that our monitoring approach could be applied to modestly sized ISP networks, and that larger networks could be monitored using parallel processes, where each process would be responsible for monitoring a particular subset of the network. Other researchers have also investigated, with promising results, techniques to characterize network-wide traffic by observing carefully selected measurement points. Chua and colleagues [24-25], for example, applied linear algebraic methods to select specific subsets of three to nine network paths to measure in order to adequately characterize time-varying averages, totals and differences in network-wide delays.

In summary, we have developed at least one technique that appears capable of monitoring a modestly sized ISP network, and we conceived an approach to extend the scope of that technique to larger ISP networks. Further, we demonstrated that our technique is capable of detecting shifting spatiotemporal traffic patterns associated with stealthy distributed denial of service attacks. Other researchers have also investigated techniques to characterize network-wide properties from relatively small samples. For these reasons, we are confident that we can devise detection algorithms capable of monitoring any system at the size of a single ISP and smaller, and also that such scope of monitoring can reveal useful pattern shifts. Monitoring the entire global Internet is impractical and unnecessary.

References:

[1] M. Takayasu, H. Takayasu and K. Fukuda, “Dynamic phase transition observed in Internet traffic flow”, *Physica A* 277 (2000) 248-255.

[2] P. Echenique, J. Gomez-Gardenes and Y. Moreno, “Dynamics of jamming transitions in complex networks”, *Europhys. Lett.* 71 (2005) 325.

[3] D. De Martino, L. Dall’Asta, G. Bianconi and M. Marsili, “Congestion phenomena on complex networks” *Physical Review E* 79 015101 (R) (2009).

[4] J. Yuan and K. Mills, "Monitoring the Macroscopic Effect of DDoS Flooding Attacks", *IEEE Transactions on Dependable and Secure Computing*, 2/4, October-December 2005, pp. 324-335.

[5] E. Coffman, Z. Ge, V. Mishra and D. Towsley, "Network Resilience: Exploring Cascading Failures within BGP", *Proceedings of the 40th Annual Allerton Conference on Communications*, 2002.
[6] H. Liao, J. Apt and S. Talukdar, "Phase Transition in the Probability of Cascading Failures", Carnegie University Technical Report, December 2004.
[7] B. Carreras, V. Lynch, I. Dobson and D. Newman, "Critical points and transitions in a electric power transmission model for cascading failure blackouts", *Chaos*, 12(4), (2005) 985-994.
[8] M. Scheffer et al., "Early warning signals for critical transitions", *Nature*, v. 461, no. 3, Sep. 3, 2009, pp. 53-59.
[9] J. Broussard, NANOG thread on Contention/Oversubscription maths, May 27, 2011.
[10] http://www.nanog.org/
[11] http://www.ietf.org/
[12] J. Haines, D. Fried, J. Korba and K. Das, "The 1999 DARPA off-line intrusion detection evaluation", *Computer Networks*, 34 (2000), 579-595.
[13] The Internet 2 Observatory Data Collections, http://www.internet2.edu/observatory/archive/data-collections.html
[14] Protected Repository for the Defense of Infrastructure against Cyber Threats, https://www.predict.org/Default.aspx?tabid=40
[15] The Cooperative Association for Internet Data Analysis (CAIDA) Data Sets, http://www.caida.org/data/overview/
[16] RIPE Network Coordination Centre Data Sets, http://labs.ripe.net/datarepository/data-sets
[17] K. Heidemann and C. Papdopoulos, "Uses and Challenges for Network Datasets", *Proceedings of Conference for Homeland Security*, 2009, 73-82.
[18] University of Illinois Urbana-Champaign Repository of Availability Traces http://www.cs.uiuc.edu/homes/pbg/availability/
[19] Network Economics Group Data Sets, http://netecon_group.tmit.bme.hu/source-codes
[20] University of California Riverside Archive, http://networks.cs.ucr.edu/ucrchive/measurement.htm
[21] Google M-Lab Data Sets, http://www.measurementlab.net/data
[22] Technische Universiteit Delft, Peer-to-Peer Trace Archive, http://p2pta.ewi.tudelft.nl/pmwiki/?n=Main.Home
[23] K. Mills, E. Schwartz and J. Yuan, "How to Model a TCP/IP Network using only 20 Parameters", *Proceedings of the 2010 Winter Simulation Conference* (WSC 2010), Dec. 5-8, Baltimore, MD, pp. 849-860.
[24] D. Chua, E. Kolaczyk and M. Crovella, "Efficient Monitoring of End-to-End Network Properties", *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Society*, 3 (2005), 1701-1711.
[25] D. Chua, E. Kolaczyk and M. Crovella, "Network Kriging", *IEEE Journal on Selected Areas in Communications*, 24(12) (2006), 2263-2272.
[26] L. Fisher, Crashes, Crises, and Calamities: How We Can Use Science to Read the Early-Warning Signs, Basic Books, New York (2011), 256 pages.

[27] Y. Moreno, R. Pastor-Satorras and A. Vespignani, “Epidemic outbreaks in complex hetergeonous networks”, *The European Physical Journal B - Condensed Matter and Complex Systems* 26(4) (2002), 521-529.

[28] Y. Tang and S. Chen, “Defending against Internet worms: a signature-based approach”, *Proceedings of the 24th Annual Joint Conference of the IEEE Computer and Communications Society*, 2 (2005), 1384-1394.