

Monitoring and Controlling the Nation's Critical Infrastructure

Introduction:

This report summarizes the conclusions of the National Science Foundation Workshop on Monitoring and Controlling the Nation's Critical Infrastructures. The workshop was organized by NSF Program Directors Dr. John Cozzens and Dr. Sankar Basu and was held on November 17-18, 2006 at the NSF. The workshop participants were Ian Dobson (UW-Madison), Maria Ilic (CMU), Pierre Moulin (UIUC), Jose Moura (CMU), Rob Nowak (UW-Madison), Kannan Ramchandran (Berkeley), John Tsitsiklis (MIT), Jeanne VanBriesen (CMU), Venu Veeravalli (UIUC), and Alan Willsky (MIT).

The increasing availability of diverse sensing mechanisms, wireless communications, and ubiquitous computational power offer unprecedented opportunities for monitoring and controlling the infrastructures that support our energy, utility, information, utility, healthcare and transportation systems. Despite the fact that these infrastructures have disparate origins, they are all examples of distributed networked systems, and often have a largely decentralized organization. Consequently, common engineering and mathematical challenges may arise in the analysis, monitoring and optimization of these infrastructures. The need for intelligent information processing in Critical Infrastructures is clearly evident. Signal processing theory and methods aimed at distributed and decentralized monitoring and decision-making, learning and understanding spatiotemporal patterns of behavior in large networked systems, identifying network structures and inferring pathways of information flow, and controlling distributed systems are crucial to optimizing the performance and preserving the operational integrity and security of the Critical Infrastructures. Some of the key opportunities will integrate signal processing approaches with modeling and analysis of specific infrastructures.

Critical Infrastructures

Critical Infrastructures of particular interest to the Program are listed below, with a brief description of the challenges posed in each case.

- Power Grid– Blackouts are catastrophic events with significant costs to society. The de-regulation of energy markets poses significant new challenges.
- Energy Networks– Gas and oil pipelines stretch over vast geographic areas and are vulnerable to failures and disruptions. Distributed monitoring and control mechanisms are required to secure these systems.
- Internet– Malicious activity poses an extreme threat to our information infrastructure, and distributed measurement and automatic response mechanisms are needed to combat this threat.
- Water Distribution– Water distribution networks are largely open systems, and distributed sensing is needed to monitor quality and regulate distribution.

- Environmental and Agricultural Systems– Detecting environmental threats or optimizing agricultural practices calls for spatiotemporal monitoring and control over large areas.
- Biological Systems– Combating disease and man-made biohazards requires a fundamental understanding of biological systems and cellular/genomic functionality. Knowledge of biological networks is very incomplete at this time.
- Surveillance Networks– Critical to the nation’s security, surveillance networks require the assimilation and fusion of information from a vast diversity of sensors and information sources.
- Transportation Systems– Intelligent transportation systems rely critically on distributed sensing and control mechanisms.

While technological solutions need to be carefully tailored to the requirements of specific infrastructures, there are some common themes to the challenges they present.. Local information gathered by a large number of heterogeneous, multimodal, possibly dynamic, and un-calibrated sensors needs to be deployed locally to further global goals and/or be fused for global decision-making. In some scenarios the geometry of sensing may be unknown or only partially known, and the environment may typically be complex and dynamic. Here a critical element of the information processing challenge is the design of effective learning tools, analysis of patterns, and adaptive strategies to optimize decision-making process. Detection of localized anomalous behavior, understanding how failures propagate in systems of large-scale networks and to cause catastrophic results, investigating practical approaches to mitigate these effects by recognizing evolving patterns of local/global behavior are major themes of the program.

Numerous challenges are presented by Critical Infrastructures, including

1. System dynamics may be largely unknown, complex and difficult to predict, especially due to the fact that systems are continually evolving themselves.
2. Complexity and dynamics are manifest over very large ranges of scales (e.g., nanoseconds to years, nanometers to kilometers).
3. Infrastructures are inherently distributed and decentralized, with many stimuli, influences and actors that are difficult to model.
4. Complete information and awareness of the infrastructure is often lacking.
5. Measurement, data-collection and experimentation capabilities may be very limited and/or costly.
6. Legacy infrastructures are being utilized in new and untested ways.

These challenges lead to inefficient performance, inadequate capacity for increasing demands, fragile/vulnerable systems, and potentially catastrophic events. The challenges may be addressed in many ways, including systematic methods for incorporating highly flexible just in time and just in place improvements, radical new methods to augment and enhance existing infrastructures, and completely revolutionary designs of future infrastructures. Monitoring and control should adapt automatically to detect and manage the most critical events, and should be capable of handling unexpected, emerging behaviors. In many instances, there will be human operators in the loop, and new

methods are necessary to help humans sort through the vast informational clutter inherent in Critical Infrastructures and better understand how humans interact with the infrastructures.

Research Elements of the Critical Infrastructures Program

1. Measurement, Monitoring and Communications: Wireless communications and the emergence of sensor network technologies have enabled fundamentally new sensing architectures capable of monitoring large spatially distributed phenomena. Research into new communication protocols, fundamental limits and trade-offs between communications and other basic system resources such as power, and optimal deployment strategies are needed. The phenomena to be measured or monitored need to be understood so that meaningful signals are processed and result in information that can be used.

2. Modeling: Modeling complex networked systems requires domain expertise and is facilitated by dynamic systems analysis, machine learning and system identification, dimensionality reduction, multiscale (spatial and temporal) analysis, distributed systems and network theory, stochastic processes, Monte Carlo methods, and statistical inference. New methodologies are required for blending top-down (physics-based) modeling with bottom-up (statistical analysis and machine learning) modeling and with simulation-based methods to yield models for phenomena at multiple granularities. Theory and methods are needed to devise models that are “focused” or aimed at particular objectives (e.g., anomaly detection, fault localization, pattern recognition, network identification). Methodologies for producing structured models for scalable estimation and control, and for capturing rare anomaly propagation behavior are especially relevant.

3. Fusion: Data and information fusion is a central task in critical infrastructure monitoring and control. Basic tools for this task include statistical inference, machine learning, detection theory, probabilistic graphical models and other distributed/networked models, combined with risk-sensitive decision theory and database search and information retrieval methods. Challenges arising in Critical Infrastructures include the robust detection of weak and distributed signals embedded in complex and highly variable backgrounds, scalable inference algorithms that exploit structure across space, time, network and scale, and rapid detection and pattern recognition algorithms.

4. Resource Management: Systems for efficient data collection, experimentation design, dynamic planning, information search and retrieval, and communications are crucial to the development of large-scale systems. New research is necessary to devise methodologies for identifying data needed for particular modeling objectives (granularity, focus) for complex distributed systems, efficient methods for real-time confidence

estimation for quickest detection, effective methods for dynamic and adaptive management of large numbers of distributed sensors, optimal resource allocation (power, bandwidth) and efficient database search methods to support anomaly detection and monitoring.

5. Interacting with Human Operators: Traditionally, human operators have largely managed many Critical Infrastructures, and these experts possess a wealth of experience and skills necessary to this task. However, as these systems become more complex and decentralized, the overwhelming volume of data and information presented to them is significantly challenging human operators. These facts present new challenges and opportunities. Automatic methods are needed to learn aspects of system behavior, potential system vulnerabilities, and human objectives through interaction of inference algorithms and human operators. Especially important is the need to recognize the capabilities and limitations of human operators, and automated methods that produce understandable outputs that help operators cope with “information overload” are very crucial. This includes semi-automatic algorithms and reinforcement learning methods that might query a human operator to resolve ambiguous hypotheses or that present the human with a manageable set of alternatives for his/her attention.

6. Decentralized and Distributed Systems: Due to the large and complex nature of Critical Infrastructures, decentralized and distributed methods for monitoring and control are envisioned as practical and scalable solutions. Theory and methods from dynamic game and team theory, decentralized estimation and detection, distributed and graph-based algorithms, and stochastic control are particularly relevant in this regard. Important directions for new research include the development of scalable algorithms for distributed monitoring and control, performance guarantees for decentralized fusion algorithms, and methods for team-based fusion involving information push and pull and objectives push and pull. Distributed control strategies may need to account for the impact of control actions not only on system behavior but also the enhancement or obscuration of the measurable effects of the anomaly to which response has been initiated.

7. Adversarial Modeling: Malicious activities are a real threat that must be anticipated in Critical Infrastructures. Game-theoretic methods and adversarial models are important tools in this regard. New methodologies for identifying system vulnerabilities, determining risks under varying conditions of adversarial capabilities and knowledge, and assessing the performance of monitoring and control algorithms in the presence of “designed” errors and attacks is of fundamental interest to the Program.

9. Security: The incorporation of more sophisticated monitoring and control mechanisms also introduces additional potential vulnerabilities to attack and infiltration and additional failure modes. New research into methods for assessing the impact of informational attacks on distributed monitoring, decision-making and control, and methods for

achieving secure information fusion and tradeoffs among complexity, reliability, security, and performance is extremely important.

Grand Challenges

Theoretical Grand Challenges:

1. Revolutionary theory and methods for reliable and predictable closed-loop sensing and control in large-scale distributed systems.
2. Distributed reliability and dynamic re-organization of networks to optimize survivability and integrity.
3. Data-driven adaptation of models and analysis to optimize different performance objectives through sensing and actuation
4. Fundamental limits of performance and control
5. Near real-time systems for distributed monitoring and control.

Concrete Example Challenges: By 2012,

1. Increase efficiency of large-scale *power grid* by 10%
2. Implement sensing and actuation into the *power grid* to reduce risks of large catastrophic failures (frequency of occurrence and number of affected users reduced by an order of magnitude).
3. Implement sensing/actuation to facilitate penetration of distributed energy resources (*renewable energy*).
4. Devise a distributed and decentralized system for detecting contamination (e.g., cholera) in drinking *water distribution systems* and for taking active responses (e.g., automatic and targeted chlorination).
5. Deploy an operational a system that can (autonomously and rapidly) detect distributed and emergent threats such as previously unseen *Internet viruses* and *biochemical hazards*.
6. Develop computational methods capable of automatically reconstructing *biological signaling networks* from high-throughput experimental data sources.
7. Practical monitoring and processing that quantifies risk of catastrophic failure of large interconnected infrastructure networks.