NITRD LSN Workshop Report on Complex Engineered Networks

September 20-21, 2012
Washington, DC

Sponsored by

Air Force Office of Scientific Research	**Department of Energy**	**National Science Foundation**

NITRD LSN Workshop Report on Complex Engineered Networks

September 20-21, 2012
Washington, DC

NITRD LSN Workshop Committee

Workshop Co-Chairs:

- Mung Chiang, Princeton University
- Nagi Rao, Oak Ridge National Laboratory

Steering Committee:

- Bob Bonneau, Air Force Office of Scientific Research
- Bryan Lyles, National Science Foundation
- Thomas Ndousse-Fetter, Department of Energy
- Sandy Landsberg, Department of Energy

Breakout Session Leads:

- Internet: Bob Doverspike (AT&T Labs) and Keith Ross (NYU Polytechnic)
- Wireless Networks: Edward Knightly (Rice University) and Stuart Wagner (Applied Communication Sciences)
- Cyber-Physical Networks: Massoud Amin (University of Minnesota) and Adam Drobot (Open Tech Networks, formerly Telcordia)

Table of Contents

NITRD Complex Engineered Networks Workshop Report

Executive Summary

Complex engineered networks are everywhere: power grids, Internet, transportation networks, and more. They are being used more than ever before, and yet our understanding of them remains limited. These networks have evolved into complex systems with behaviors and characteristics that are beyond the characterizations and predictions possible by the traditional modeling, analysis and design approaches.

This workshop brought together experts from academia, national laboratories, government, and industry to assess the recent trends, state-of-the-art research, and impending challenges in modeling, predicting, and controlling the behaviors of these complex engineered networks.

To gather strong motivating examples, one needs to look no further than recent headline news, ranging from the significant power outages following Hurricane Sandy and the multi-day service outage of Amazon clouds, to the mysterious under-performance of mobile services and the lack of holistic privacy protection on the web. Motivated by these critical needs in this scientific community, this workshop is oriented around the following three "umbrella terms" of the charge statement:

1. What are the "big questions" in complex engineered networks research?
2. What are the grand challenges in "methods" for analysis, design, deployment, and operation of these networks?
3. What are the mechanisms to "ensure the highest impact" of federal research support?

We identified five areas where progress has been made and much more science is sorely needed: 1. revisiting architecture and abstractions, such as layering principles counting sessions rather than bytes, 2. sharpening our understanding of the limits of performance by unifying theories like those of Shannon and Bode, 3. quantifying the principles of evolvability of networks, such as loose coupling and overlay, 4. achieving robustness to dynamics, attacks and mismanagement, and 5. understanding transient behavior over time as well as we understand equilibrium behavior today. Collectively we call this research agenda ALERT: Architecture/Abstraction, Limits, Evolvability, Robustness, and Time.

Part I
Overview

Workshop Background

Complex engineered networks are everywhere – power grids, Internet, transportation networks, and others. They are being used more than ever before, and yet our understanding of them remains limited. The Internet, wireless networks, and online social networks have shaped modern society. Increasingly, critical, engineered, large-scale systems, such as transportation networks, power grids, and wireless communication systems, are being enhanced and optimized by state monitoring and dynamic controls through sensor and cyber mechanisms. These networks have evolved into complex systems with behaviors and characteristics that are beyond the characterizations and predictions possible by traditional modeling, analysis and design approaches.

This workshop will bring together experts from academia, national laboratories, government, and industries to assess the recent trends, state-of-the-art research, and impending challenges in modeling, predicting and controlling the behaviors of these complex networks to gain better performance, efficiency and robustness. The objectives of the workshop include:

- Identify transformative research challenges and directions in the field of large-scale, complex engineered networks and interconnected physical systems of sensors and instruments, such as the power grid and communications networks.
- Assess the state-of-the-art research, future trends, and important opportunities and challenges in the theory, design, analysis, tools, and applications of complex interconnected systems research in government, industry, and academia.
- Identify strategies at the federal funding agencies to enable different communities to carry out joint efforts, leverage ongoing activities, accelerate new discoveries, and enable technology transfers for societal impact in the research field of complex networks and interconnected systems.

What is a Complex Engineered Network?

A network consists of potentially disparate nodes, such as user devices, sensors, switches, control centers, computing and data servers, which may be distributed in both physical and cyber space. It is characterized by its topology (nodes connected by links) and a set of functionalities running on top of that topology. Engineered networks are complex systems

designed by and operated according to engineering processes by human beings, with specific objectives (such as performance and robustness) and constraints (such as physical or economic limits). Complex engineered networks are those engineered networks where the topology is neither completely regular nor completely random, or where functionalities have poorly understood behaviors. In other words,

- A network consists of not just a collection of nodes and links (which would constitute a graph), but also an evolving set of functionalities, including operation protocols, control, and deployment.
- An engineered network is a network designed and operated by human beings for predictable performance, thus bringing the human factor as a potential root cause for complexity.
- A complex engineered network is an engineered network with properties that are hard to understand or predict (sometimes though not always when it is stretched in physical scale or temporal evolution).

An engineered network is one that maybe complex, not completely understood, but we still have the expectation that it will behave predictably under a wide set of circumstances. Much of the effort over time is to better understand the behavior of operating networks and their classification, and to improve the predictability of their attributes.The overarching questions on the science of complex engineered networks include: Can an evolving network be designed to be less complex over time? Can there be a science of network management? Can there be scalable tools for the design and operation of such networks?

Examples of Complex Engineered Networks

Complex engineered networks are playing increasingly critical roles in our society, and their role is expected to continue to expand both in scale and scope. They span a wide spectrum, encompassing the Internet and multiple modalities of wireless networks, power grids with smart homes and cars, green house gas monitoring networks with satellites, airborne and ground sensors across the world, global networks of telescopes, and networks of instruments and sensors from battlefields to hospitals. We now describe three specific examples that represent three different types of network and application areas within the vast space of complex engineered networks.

Example A: Long-haul sensor and communication networks: Sensors are distributed across the globe and/or in space to support tasks such as (i) monitoring of greenhouse gas emissions using satellite, airborne, ground and sea sensors; (ii) processing global cyber events using cyber sensors over the Internet; (iii) space exploration using a network of telescopes on different continents; and (iv) target detection and tracking for air and missile defense. The response

time requirements are quite varied, ranging from seconds for detecting cyber attacks on critical infrastructures to years in detecting global trends in greenhouse gas emissions. These complex networks are quite different from the well-studied "smaller" sensor networks since the sensor controls and information flows must be coordinated over connections that could range over several tens of thousands of miles.

Example B: Detection of low-level radiation sources: Wireless Networks of detectors are deployed to detect, localize and track sources based on sensor measurements in a wide range of scenarios, ranging from events at sports stadiums to border crossings to national highways to ship yards. The deployments could be ad hoc wireless networks of detectors at events or highway detector weigh-stations connected to interstate monitoring control centers. The response times could vary from a few seconds in the first case to detect potential sources, to days to track down long-haul movements of certain materials.

Example C: Experimental network research testbeds: To support the design, deployment and testing of next generation and advanced networking and related application technologies, experimental network facilities are built and operated. For example, GENI infrastructure supports open research projects, while ESnet DSN and UltraScience Net support high performance networking projects for specific agency needs. The requirements of such facilities are often quite different from the provider networks, in that they enable users to experiment with structural reconfigurations, whereas structural stability is of paramount importance in traditional networks.

During part of the workshop, we divide the discussion into three types of complex engineered networks: (i) Internet, (ii) wireless networks and (iii) cyber-physical networks. Parallel break-out sessions are onducted in each of these areas, and their brief summaries are included as appendices.. The research opportunities and overall directions are distilled to cut across these areas and presented in the next sections.

Research Opportunities

Our focus includes design, implementation, testing, deployment and operation to ensure robust, optimized and cost-effective performance. These networked systems promise capabilities unprecedented in their performance, scale, and applications. But their sheer complexity and scale makes it very challenging to understand and predict their performance across a full set of operating conditions. Formal mathematical analysis may give insight into the operation of system components or into the operation of simplified system models, but it is seldom capable of guiding the overall designs and physical realizations. Empirical testing of deployed systems, on the other hand, often misses unexpected behaviors resulting from cross-system interactions in these inherently infinite-dimensional systems, due to the limitations of measurements and computations. Systems engineering may provide an overall approach to

deal with the issues in complex engineered networks at an abstract level, but it cannot be reduced to practice without specific tools that yield qualitative and quantitative results to guide engineering principles. These tools are domain specific, span many disciplines, and yet must operate in a coherent framework to have practical impact. Such a framework does not exist today, thus presenting a significant opportunity to advance the science of complex engineered networks.

To gather strong motivating examples for the sore need for more science in this field, one needs to look no further than recent headline news in the media, ranging from the significant power outages following Hurricane Sandy and the multi-day service outage of Amazon clouds, to the mysterious under-performance of mobile services and the lack of holistic privacy protection on the web. Motivated by these critical needs in this scientific community, this workshop is oriented around the following three "umbrella terms" of the charge statement:

1. What are the "big questions" in complex engineered networks research?
2. What are the grand challenges in "methods" for analysis, design, deployment and operation of these networks?
3. What are the mechanisms to "ensure the highest impact" of federal research support?

Part II
Research Opportunities: ALERT

A for Abstraction / Architecture

At the heart of complex engineered networks are questions about architectures. An architecture is a scheme of functionality allocation: what should each module do, at what timescale, and how do we glue them back together? This is often a much less understood and yet more influential design decision than decisions about resource allocation, such as which paths to take in a network or how to allocate frequency bands to competing radios. As complex engineered networks increasingly take the form of a system of inter-dependent networks, such as mobile social networks and smart grid power networks, architectural decisions are even more challenging to make across modules, across data plane and control plane, and across end users and the rest of the network.

To understand architectures is to understand modularization. In control systems, serial computation systems, and point-to-point communication systems, architectures have been well-studied with an abstraction that is rigorous and relevant at the same time, thanks to the architectural foundations present in control theory, theory of computation, and information theory. Not so, yet, for complex engineered networks. In control systems, there is the plant-observer-controller-actuator division of labor. In computing systems, there is the input-processor-memory-output division of labor. In communication systems, there is the source-channel separation principle. However, a top-down, first-principled design "language" to architect the layering of complex engineered networks remains elusive.

In order to study architectures, we first need to develop models that explain more than just describe, to derive conclusions that provide predictions in addition to hindsight, and to build theories that can be falsified by empirical data rather than merely self-consistent. For example, the mechanics of congestion control in Internet's Transmission Control Protocol (TCP) [FS11] itself is a descriptive model, but the reverse engineering of the protocol into an implicit solution to a capacity-sharing optimization problem provides an explanatory model [KMT98]. Such explanatory models provide insights on when and why certain engineering artifacts work the way they do in a complex system. They sometimes also possess predictive power as they guide forward-engineering, such as the design of highly efficient and stable congestion control methods [LDP02]. While assumptions are made in both the modeling and design process, the theory makes predictions that can be falsified in experiments.

Architecture design presumes the objectives of a complex engineered network. Some objectives have taken up most of our attention, often because they are more tractable, even though other

objectives may be "higher-order bits" in the actual operation of a complex network. For example, our understanding of throughput performance is significantly better than that of delay properties, which is in turn much better than that of security and privacy metrics, which is in turn much better than that of manageability of the network. To properly characterize delays, let alone optimize for delay performance, we often have to make the choice between the microscopic abstraction, where discrete events to each packet are characterized along a path in a network, and the macroscopic abstraction, where a fluid approximation is taken. Treating packets as fluids simplifies the queueing-theoretic challenges, but constrains us to questions that can only lead to approximations of actual delay performance. Furthermore, while it might be difficult to optimize for delay, it is not even clear what metrics should be used to quantify certain aspects of security and privacy, or even how to define the manageability of a complex network. Network operators can provide many anecdotes of poorly managed networks and network design with poor manageability. Except for recent work using variations of formal methods to certain dimensions of enterprise network management [BAM09], we do not yet have a theory to bring the manageability issues to a sharp focus for debates.

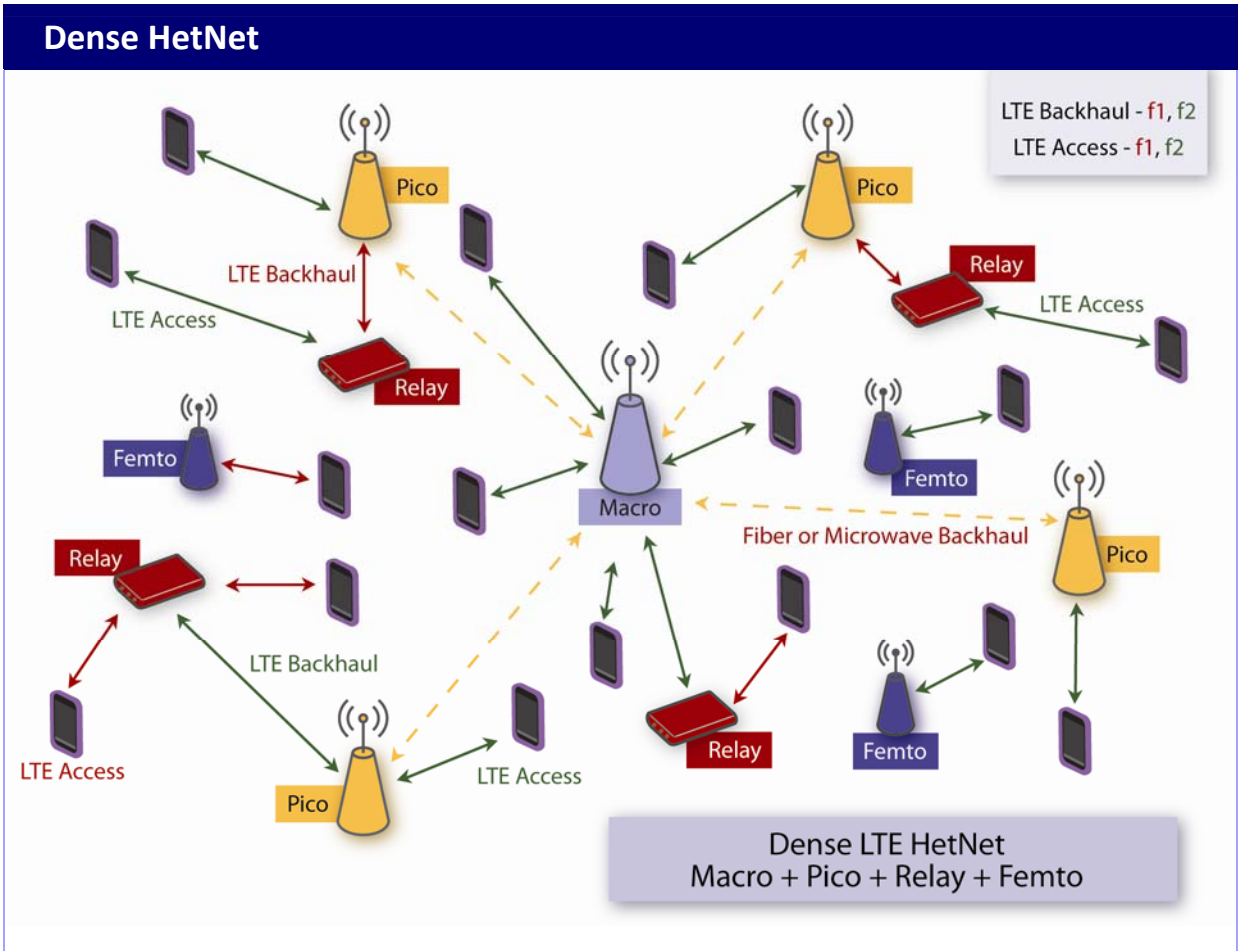

Figure 1: Wireless networks heading to dense deployment of heterogeneous networks with the coexistence of macros, picos, femtos, WiFi, and relays

In general, it remains challenging to quantify objectives to maximize beyond the traditional performance metrics. The "network x-ities," e.g., evolvability, manageability, scalability, and diagnosability, these somewhat fuzzy English terms, are critical in network operations and situate right in the core of understanding the complexity of networks. Yet they are often the least well-understood and much less studied, compared to performance metrics such as throughput, delay, energy, distortion, and jitter. Some of these x-ities will be discussed further in later sections. Similarly, fairness as a metric is elusive to quantify, characterize, and optimize for. Multiple disciplines have all studied quantification of fairness, and recent results point to a unification of these fairness metrics in a set of simple axioms.

Complexity itself can be a key component in the objective function for network optimization, and not just computational complexity or communication complexity, both of which are quite well-studied, but also configuration complexity, e.g., the number of the tunable parameters and the effort it takes to tune them to the right range of values for robust performance. It makes performance optimization easy on paper by adding more and more tunable parameters that adapt to observations. But in real network operations, observations are often noisy and delayed, and adaptation rules unclear as the environment changes rapidly. Even tuning the parameter in the training phase can be difficult. Sometimes parameter tuning, such as in Random Early Detection [FJ93] for buffer management in the Internet, leaves network operators hesitant to turn on the feature. At this point, there are so many variants of control protocols and so many tunable parameters in each variant that it has become almost impossible to understand the mutual interactions of existing parameters, or to predict the global effect if an additional parameter is introduced. In attempting to operate a network more efficiently, man-made configuration complexity rises rapidly. Often in the process of achieving better performance, engineers make the network's inner workings and observable behaviors more complex. This readily leads to a vicious cycle: more protocols and parameters are piled upon the complex network and make it even more complex.

It is essential to have the "right" level of abstraction, and select the corresponding level of approximations, not just based on the convenience of tractability but on an overall approach that selects the critical components of each abstraction. The packet vs. fluid debate mentioned above is a typical example. Some more recent examples include:

- "Approximate networks," where information is sampled and packets delivered in approximate way by design; e.g., compressed sensing [CRT06, D06] and network coding [ACLY00, KM03].
- Sufficient statistics of graphs so that we can carry out operations on these complicated, discrete objectives through simpler representations.
- Aggregation of dynamics on graphs, including the dynamics of cascading failure [B+10] and propagation of influence [M00].

Related to the level of abstraction to operate our research, we need to determine the fundamental logical unit in a network: is it a session or a piece of content? Counting bytes or measuring transactions? Research in content-centric networking (or, "named data networking") [J+12] has created new angles to look into the management of states from an end-user point of view, hoping to restructure layering, naming, and security in communication networks. Recent work on smart data-pricing also points toward billing and charging based on each transaction's value to end users, rather than a simple counting of the bytes that pass through a point on the data plane [SJHC13].

We need to carefully manage the types of states and awareness to be maintained within and outside a complex network. The debate of soft state vs. hard state [LMR04] has run its course for over two decades, and complex networks designed through industry standardization bodies have made many decisions as to which network element maintains what kinds of state, and which measurement counters of the network or its external environment should be exposed to each control knob. What is in short supply, however, is a systematic design procedure for choosing among the alternatives of these decisions.

Across all of the above challenges in understanding the abstractions and architectures of a complex engineered network, we need a feedback loop in the research process: from models to theory, then to design, then to experiments, then to deployment, then to empirical data, and back to models. Scalable, realistic, and yet cost-efficient experiments need to be designed, and at-scale experiments must be carried out to validate or falsify the theory of complex networks. Some of these challenges will be addressed in Part III of this report.

Nugget Case Study: Distributed Architecture in Smart Grids.

From its inception, the power grid has been susceptible to rolling blackouts in many areas. This is especially the case in countries that have a centralized distribution structure, where power is generated from a small number of large plants and feeds out to customers in a hierarchical manner. Here, the failure of a single plant could cause outages for tens of thousands of homes, and even a single feeder line can be responsible for the flow of power to many customers. We have seen many examples of this in recent history.

In August 2005, Hurricane Katrina caused widespread power outages throughout Louisiana, Mississippi, Alabama, Florida, Kentucky and Tennessee. Power had also been also knocked out to 1.3 million customers when Katrina passed over Florida several days earlier. Similarly, the July 2012 India blackout was the largest power outage in history, occurring as two separate events on 30 and 31 July 2012. The outage affected over 620 million people, about 9% of the world population, or half of India's population, spread across 22 states in Northern, Eastern, and Northeast India. This event was triggered by a failure in a single line (the 400 kV Bina-Gwalior line). More recently, Hurricane Sandy swept through the Eastern Seaboard of the U.S. in October 2012, leaving over 5 million homes without power, some for a week or more.

Part of the problem is that the failure of a single line serving customers without a backup path will leave them in the dark until the problem can be fixed. Sometimes, simply locating and diagnosing the problem can take a significant amount of time. As a result, power companies have been advocating two potential solutions: (1) distributed generation, and (2) redundant distribution structures. Having a large number of smaller generators will make the power sources more local, and then naturally the failure of any given source will have less of an effect on the entire system. Adding redundancy implies that when a feeder path fails, there is at least one other way for power to flow into the area in question, thereby preventing a partition in the grid. As another possibility, rather than distributed sources, companies could add significant storage capacity near homes, allowing them to keep their electricity for a given amount of time before repairs are completed. PSE&G in particular has recently announced that they will launch a massive effort to integrate localized, renewable forms of energy into the grid to mitigate redundancy issues.

Nugget Case Study: Information Spread in Online Social Networks

Online social networks like Facebook and Twitter are reshaping the way people take collective actions. They have played a crucial role in the recent uprisings of the Arab Spring and the London riots. It has been argued that the 'instantaneous nature' of these networks influenced the speed at which the events were unfolding. Another popular example of this phenomenon is the viral spread of the song "Gangnam Style" by PSY. While most of its publicity is measured through YouTube likes, social networks have played a crucial role. A study shows that YouTube had around 2,072,665 PSY subscribers out of 800,000,000 active users, which is a percentage of 0.26%. Facebook has seen 4,125,318 PSY Likes out of 1,000,000,000 active users (0.41%) and Twitter has 1,649,836 PSY followers among 150,000,000 active users (1.10%).

It is quite remarkable that social networks spread news and fads so fast. Both the structure of social networks and the process that distributes the news are not designed with this purpose in mind. Is our view correct that social networks ease the spread of information, and if so, what particular properties of social networks are the reasons for this? To answer these questions, researchers have tried simulating rumor spreading processes on several graphs created with the structure of existing large social networks. Further, they attempt to define bounds and limits on how fast this information can spread, subject to the structure of the underlying interaction graph, the initial nodes that were picked and the nature of information exchange.

L for Limits

There is still much to be done in characterizing the boundary of performance in complex engineered networks, and even more challenges in understanding the boundary of unknown behaviors. This challenge is becoming more acute as complex engineered networks become a convergence of communication, computation, and storage, each of which has been extensively studied in the past with its own set of metrics. In particular, the set of theories which is built on largely ignoring the delay dimension, such as information theory [CT06], and the set of theories which is built on focusing on delay, such as control theory [AM08], remain to be unified to provide a language about the boundary of performance. Some of these theories, such as queueing theory, have also been faced with very difficult open problems when situated in a network with a general topology [K75], to the point that it is worth asking whether more tractable problems should be raised instead. This echoes the question of abstraction (fluid model or packet model) in the last section.

Besides the modeling languages of information theory, control theory, and queueing theory, optimization has also become an often-used approach in the analysis and design of complex engineered networks. The watershed between easy and difficult optimization problems has long

been recognized as “convexity” [BV04]. Non-convex optimization appears in many analysis and design problems in complex engineered networks, often with suboptimal local optima where the design may be trapped. Finding efficient and distributed solutions with performance guarantees to non-convex optimization is in general extremely difficult, unless there is a substantial amount of structures in the problem, such as the Optimal Flow Problem in energy networks.

Related to that challenge of optimization is the need for distributed optimization over combinatorial structures, such as graphs [BT97], with stochastic dynamics across multiple timescales, such as coherence time, algorithmic convergence, mobility, or user behavior changes. Some of these dynamics are naturally separated by their timescale, e.g., fiber capacity placement vs. application session duration. But sometimes the timescales are coupled, e.g., channel coherence time and algorithmic convergence time. This is particularly challenging when the stochastic dynamics follow heavy tail distribution [LTWW94], as opposed to the standard assumption of Poisson arrival with exponential traffic load distribution.

Requirements on Systems and Architectures

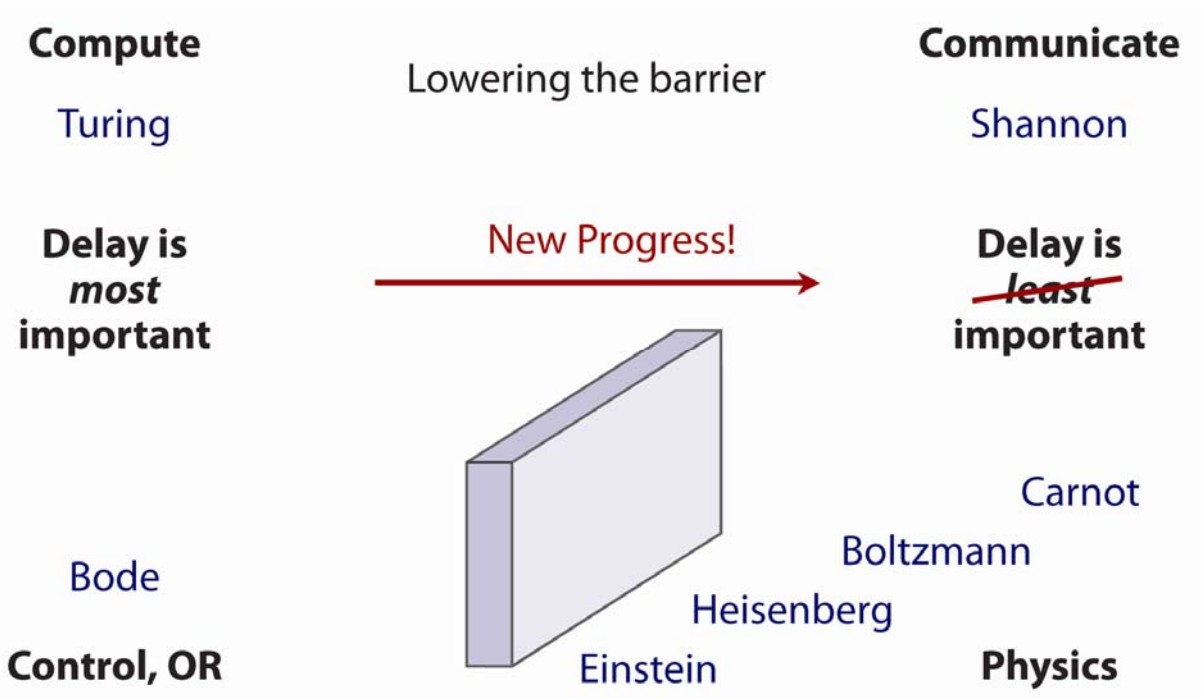

Figure 2: Diverse set of different performance limits in communication, computation, and control

Most of the results in tackling the above challenges, when they exist at all, focus on large-scale, asymptotic systems, as the number of network elements, often denoted by N, tend to infinity. There is very little mature methodology of analysis with a finite, especially small, size network population, i.e., a finite or even small N. Complex engineered networks do not have to be large; much of the complexity comes from the interactions across the functional modules and physical network elements. When a network is complex and large at the same time, it is important to track the impact of network size on the complexity [N10].

At the same time, the ability to scale up is important. In network design, scaling up the system in a cost-effective way is equally demanding, with several approaches in the design toolkit today, such as logical overlay (scaling up by building on top of existing networks [ABKM01]) and multi-stage switched networks (scaling up by scaling out [C53]). However, scaling up with respect to heterogeneous traffic demands, more than just the number of users, is more difficult and less understood. This is particularly important as the trend of moving from dedicated resource allocation to dynamic resource sharing intensifies: from circuit switching to packet switching, from dedicated server farms to rented cloud, from TDMA to CDMA, and from static spectrum allocation to tiered spectrum sharing.

Furthermore, while there are many models explaining the emergence of macroscopic behavior in a complex engineered network from microscopic interactions, there are still substantial debates on where they arise from. This is in sharp contrast to our understanding of complex networks in nature, where the transition from micro to macro behavior is often well understood through models ranging from flocking to the Ising model [V+95].

Quantifying limits in complex engineered networks is further complicated by the emergence of dynamics in systems composed of a large number of small nonlinear systems. While graph theory is often used in our study, the challenges go well beyond optimal topology design and into the growth of graphs or spread of viruses and other attacks over time, the spatial correlation and cascading faults across a graph, and the connections between graph structures and dynamics on graphs.

Data analytics interacting with network dynamics presents another under-explored area where limits of performance deserve further study, especially when the data collected are noisy, distributed, and delayed, as is typical in smart grids, wireless networks, and the Internet.

Nugget Case Study: Internet of Things

The Internet of Things was "born" in 2009, when the number of things connected to the Internet exceeded the number of people connected. By 2020, several tens of billions of devices are predicted to be connected. It is envisioned that the physical things/devices will be outfitted with different kinds of sensors and actuators and connected to the Internet via heterogeneous access networks enabled by technologies such as embedded sensing and actuating, radio frequency identification (RFID), wireless sensor networks, real-time and semantic web services, etc.

IoT is a network of networks with many unique characteristics. With the huge number of things/objects and sensors/actuators connected to the Internet, a massive and in some cases real-time data flow will be automatically produced by connected things and sensors. It is important to collect correct raw data in an efficient way, but more important is to analyze and mine the raw data to abstract more valuable information such as correlations among services.

IoT needs an appropriate architecture, such as a service-oriented, a content-centric, or a thing-centric architecture. It is also challenging to design efficient protocols to cater to diverse IoT devices, sensors and services, including many services that are very delay-sensitive rather than throughout-demanding.

E for Evolvability

The evolution of complex engineered networks constitutes a key root cause of its complexity. The evolution of networks in nature, from swarms to biochemical pathways, and from genetics to evolution of species food chains, has been extensively studied in disciplines that are devoted to these subjects. For engineered networks, the timescale of evolution has been limited to hundreds of years at most, but often on the order of a decade. This presents additional challenges in understanding the rise and fall of complex patterns and design paradigms in such networks.

The problem is even more acute as backward-compatibility, incremental-deployability, and economic incentives are key drivers of network technologies' success or failure. Many stories of success and failure in network evolution are driven by whether the technology is compatible with legacy protocols and equipment, as the timescale of phasing out a technology is much longer than that of the application needs; or they are driven by whether that technology can be deployed incrementally, with minimal disturbance and sufficient benefits even when only a fraction of the network elements are upgraded to the new technology; or driven by the business directions and financial incentives of all the key players, as exemplified by the competition between WiMAX and LTE as the dominant technology for 4th generation cellular networks.

It is often believed that the phenomenal success and accelerated evolution of the Internet is in part due to two principles. One is under-specification, including a layered protocol stack that readily allows one module to be updated without disrupting all the other modules. "APIs" are provided to enable future, unforeseen innovations through under-specified interfaces. The second principle is the ability to build overlay networks and a sequence of evolving graphs on top of each other. Still, theories of self-correcting, under-specified, and adaptive networks are yet to be built in reality. The questions on the tradeoff between the payoff (such as ease of evolution and structuring of the industry sectors) and the price (such as loss in efficiency and addition of overhead) of modular interface remain to be formulated and properly addressed.

Recently, network virtualization [P+06] has generated new opportunities to divide a given network into many "slices" for sharing. Virtualizing a network is also much more difficult than virtualizing a computer. Careful management across the control plane and the data plane is required. Virtualizing the radio air interface, or enabling agile sharing of the spectrum with sufficient isolation, remains a significant challenge.

A similar line of questions presents itself along the spatial dimension: where should centralized control be used in a complex engineered network and where should distributed control be used? The degree of "distributedness" is not a binary one; for example, one can try to quantify it by examining the frequency, spatial reach, and length of message passing needed. This also leads to the question of what level of hierarchy should be used. Often the answer varies from the control plane design to the data plane design; for example, it is common to see a somewhat centralized control plane for signaling but a more distributed data plane for actual traffic flow. This tension between centralized control and distributed control runs from the Internet (e.g., how OSPF link weights are calculated in the centralized server while hop-by-hop forwarding is done distributedly at routers [H99]) to wireless networks (e.g., how billing and charging functionalities are carried out in cellular core networks [C13] while transmit power controlling is done at the mobile devices).

Economic viewpoints of complex engineered networks interfaces with the evolution of such networks in multiple ways. For example, pricing, whether explicitly facing the users or implicitly as an angle of interpretation, has been used as a key approach to create feedback loops in complex engineered networks with a short timescale. Over a much longer timescale and more generally speaking, economic incentive mechanisms may also play a critical role in the evolution of network technologies and consumer behavior. From the economic optimization of smart grids [FERC08] to simple yet effective pricing strategies in mobile cellular networks, there is a growing need to take an inter-disciplinary view of how social, economic, and technological networks co-evolve together.

Heterogeneity is a key component of network evolution, and the complexity of engineered networks often stems from heterogeneity in traffic requirements, network platforms, and operational timescales. If bursty traffic justified packet switching, what are the implications of heterogeneity of traffic today, which includes those that can wait and those that cannot, those that can tolerate distortion and those that cannot, etc.? Heterogeneity of network platforms further adds a burden to network manageability and to user choice. Access to dense HetNets with small cells and multi-tiered spectrum-sharing presents new challenges of management complexity in wireless networks: how do we send the right bytes at the right time, at the right place, and at the right rate?

Nugget Case Study: Heterogeneity of Radio Platforms

Carriers are struggling to cope with the explosion of data traffic on their networks, and this has resulted in over 100 commercial LTE network deployments worldwide, with nearly 350 carriers committed to the 4G technology. However, the additional deployment of LTE and legacy 2G/3G RAN macro cell sites, combined with the procurement of additional frequency spectrum resources, will not be sufficient to provide the capacity for the predicted growth in traffic. Consequently, together with their on-going transition to LTE, carriers have begun to deploy Heterogeneous Networks ('HetNets') to address these ever-increasing capacity demands. HetNets are a gradual evolution of cellular topology, not a distinct network unto itself. HetNets often have three major components. The first is an umbrella or macro network designed to provide ubiquitous mobile broadband coverage. The second is a dense network of small cells that supply large quantities of bandwidth in the high-traffic areas where it is most needed. The final component is a network intelligence that ties those networks together.

Fourth generation (4G) wireless systems are devised with the vision of heterogeneity in which a mobile user/device will be able to connect to multiple wireless networks (e.g., WLAN, cellular, WMAN) simultaneously. This heterogeneity not only exists in the wireless network architecture (e.g., the co-existence of 2.5G, 3G, LTE, WiMAX, Wi-Fi, and femto, among others), but also in mobile devices (different mobile devices with different OS and functionality) and applications (e.g., low bandwidth voice vs. high bandwidth video streaming), which together enable a variety of services. Spectrum sharing and sensing over HetNets also lead to new types of complexity in wireless networks.

Nugget Case Study: Self-Healing Smart Grids and Self-Optimizing Networks

Self-healing smart grids leverage domain specific knowledge of power systems' operating states. They take into account the anticipation of disruptive events, enabling fast isolation and sectionalization, adaptive islanding, and restoration. Ideally, they will cope well with variations in their operation environments with minimal damage, alteration, or loss of functionality.

A similar trend is emerging in wireless networks. Traditionally, many network elements and associated parameters are manually configured. Planning, commissioning, configuration, integration and management of these parameters are essential for efficient and reliable network operation; however, the associated operations costs are significant. Specialized expertise must be maintained to tune these network parameters, and the existing manual process is time-consuming and error-prone. This manual tuning process results in comparatively long delays in updating values in response to the often rapidly changing network topologies and operating conditions, resulting in sub-optimal network performance.

The recent deployment of LTE to address the growing data capacity crunch has highlighted the need for and value of self-organizing capabilities within the network that permit reductions in operational expenses (OPEX). 3GPP initiated the work towards standardizing self-optimizing capabilities for LTE in Release 8 and Release 9. The standards provide network intelligence, automation and management features in order to automate the configuration and optimization of wireless networks to adapt to varying radio channel conditions, thereby lowering costs, and improving network performance and flexibility. This effort has continued in Release 10 with additional enhancements in each of the above areas and new areas allowing for inter-radio access technology operation, enhanced inter-cell interference coordination, coverage and capacity optimization, energy efficiency and minimization of operational expenses.

R for Robustness

Robustness is a highly desirable property of complex engineered networks, sometimes of first-order importance. The challenge is to construct robust network operations out of not-so-robust components, e.g., robust grids out of not-so-robust sensors. Network engineers have been pushing peak throughput higher and higher, yet robust throughput in light of changing traffic, topology, and channels can be very low. In some sense, we are still missing the dependable "dial tone" for mobile data networks. We are not close to the goal of modularly composable, real-time verifiable security, safety, and privacy on complex networks.

Robustness needs to be maintained with respect to attacks, dynamics, and mismanagement including inadvertent misconfiguration. Complexity of network management is closely connected to fragility of network operation. Offense is easier and cheaper than defense. Moreover, there are often unanticipated vulnerability and unforeseen attacks. It becomes unrealistic to have infinite resilience. We need theories that provide different levels of resilience for different users, organizations, and missions.

Optimal (with respect to certain assumptions) yet fragile (to disturbances from the assumptions) designs are often root causes of the complexity of engineered networks. There is also a subtle difference between robustness with respect to known disturbances and non-fragility with respect to unknown ones, leading to networks that are robust yet fragile.

Yet robustness is much less understood than performance metrics. Security (including confidentiality, integrity, availability, and non-repudiability) and privacy (including trustworthiness and anonymity) are clearly key elements of robustness, as is resilience (under stress, disturbances, and attacks). Some of these metrics have been well-quantified, such as availability, but the majority of them have not.

It is much easier to be resilient to known types of unknowns than to unknown types of unknowns. A major type of unknown unknowns is the disruptive jumps in technology and user behavior. These cause fundamental changes to the network design motifs and behaviors and yet by their very nature do not lend themselves to parameterized predictions. Another type is black swan events, with an extremely small probability of happening yet a disastrous amount of damage if they do happen. A naïve probabilistic characterization by mean and variance does not suffice. Quantifying resilience metrics, modeling attack modes, and reasoning about extremely rare events all present tremendous intellectual challenges beyond what the standard use of probability theory can provide today.

There are also other dimensions of robustness, especially in network management, such as diagnosability (the ability to detect the occurrence and root causes of failures) and reboot-ability (the ability to cold start the network). They represent the ability to contain and counter the damages even when they are hard to prevent in the first place. Such a science of network management is hardly existent today, even though collections of folklore theorems and practitioners' wisdom are abundant.

Robustness to our own design is just as important as robustness to the external environment. Robustness and simplicity are two axes that are often intertwined. Formal methods for invariant specification and verification in controlled network systems need to scale as fast as code complexity. Conducting experiments ideally should serve as a means to check and enhance robustness to flaws in our own design. The science of experiment design is widely used in science and engineering disciplines, but is often ignored in the study of complex engineered networks. This in turn has led to a shortage of simulations that we can believe in, of experiments driven by empirical data, and of results that are statistically illuminating and reproducible in this field.

Many of the designs in complex engineered networks exhibit the property of "optimal but fragile." [DC00] They may be optimal with respect to the stated objective, but the price for achieving such demonstrated optimality might be excessively high sensitivity to configuration parameters. A large number of adaptive configuration parameters that need to be fine-tuned (but will often face the error-prone human operations) could also result. Configuration complexity discussed in earlier sections becomes an issue of robustness. Sometimes, such design translates into complex engineered networks, especially those in the cyber-physical world, where a small action can readily lead to a cascading set of events. We can view such networks as a control dynamic system with a high "gain." Lowering the "gain" of such networks may provide a pathway to more robust design, especially in light of black swan events.

Requirements on Systems and Architectures

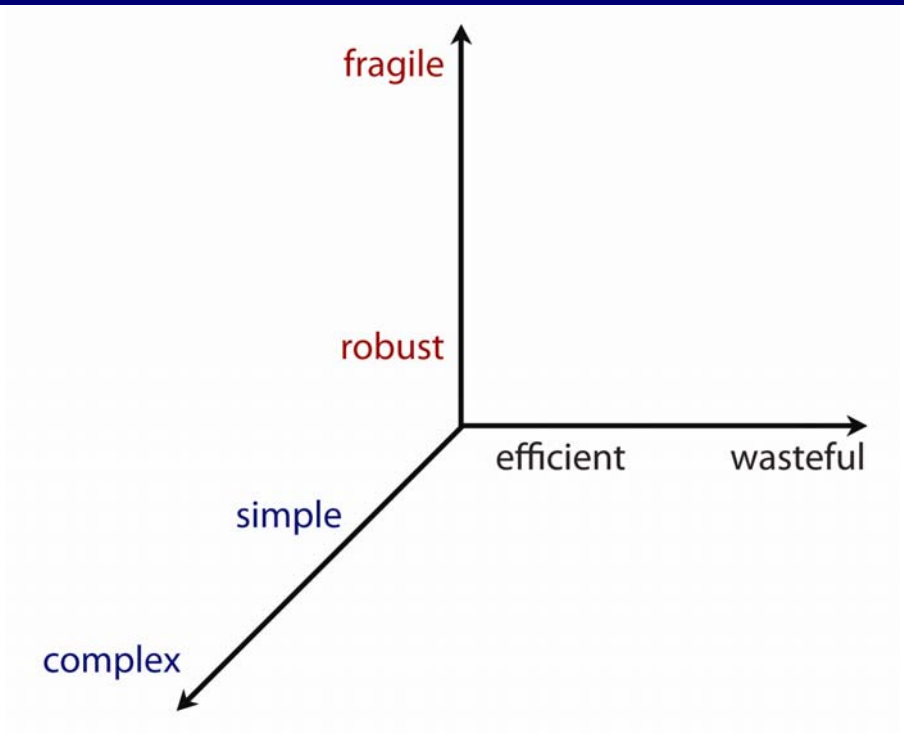

Figure 3: Tradeoff space across efficiency, robustness, and simplicity

In designing robust networks, we are also torn by two opposite forces of "diversity." On the one hand, we have to ask if there is enough diversity of resource allocation for robustness. Indeed, smart repetition and diversity are common techniques in achieving robustness in systems ranging from signal transmission to power grids. On the other hand, we would like to simplify the management of diversity, so that complexity is hidden from the end users (e.g., through automated agents or clever user-experience design like those in Apple products, Amazon kindle, or on eBay) and they are not confused by the availability of choices that can in turn lead to fragility of the network.

Robust networks also need to dynamically learn the changing environment across different timescales. Such learning may use limited computation or network resources, thus leading to a tradeoff between exploration and exploitation.

Nugget Case Study: "Safety, Security, Privacy"

Human vulnerabilities have led to many breaches. Here are some examples: (i) HBGary Federal, a beltway computer security firm, had all of its email stolen and made available on BitTorrent. (ii) The Epsilon mailing list service, which maintains mailing lists for many corporations, had its databases hacked, quite possibly through a phishing attack; (iii) The PlayStation Network was hacked, with more than 65 million accounts compromised, including names, street addresses, email addresses, and purchase histories stolen; (iv) Global Payments Inc. had a data breach that may have involved up to 1.5 million credit and debit card accounts.

Common among these attacks are two aspects. The first is that many of these attacks have shifted from just directly attacking a computer system, toward exploiting the human vulnerabilities in these systems. The human vulnerabilities include all of the misunderstandings, laxness, and cognitive and social biases that arise with the people who use computer systems. The list of human vulnerabilities here are numerous: poor interfaces that are difficult to understand, interfaces that are easy to misconfigure, guessable passwords, reused passwords, tricking people into installing malware, tricking people into opening up documents, etc.

The other aspect is that not many of these hacks are actually new or innovative. The attackers are more patient, adept at using a wide range of tools, and very capable of progressively exploiting smaller vulnerabilities to create larger ones. These recent attacks have been highly creative, flexible to the situation, and make full use of a combination of techniques. For example, in the Epsilon case, it looks like the attackers were using spear-phishing attacks for several months, trying to bait low-level employees at several mailing list companies.

Nugget Case Study: Cellular Network Robustness

Today, with the advent of cellular interfaces on their end devices such as smartphones and tablets, many consumers feel that they are now immune to voice and internet isolation due to natural disasters. However, the wireless segment of the total communication path, the portion between the end device and the cell site, is a very short piece. The rest of the communications path is dominated by landline facilities and equipment. In particular, cell sites are all connected via copper, fiber, or line-of-site wireless "backhaul" to the (wire-line) packet transport backbone of the various wireless carriers. In addition, the cell sites themselves contain a significant amount of telecommunications equipment from each carrier, such as Ethernet switches and Radio Access Network (RAN) equipment. During Hurricane Sandy there was a high loss of cellular connectivity. Some of the major reasons were 1) flooding of the cell sites, 2) damage or flooding to the landline facilities that connect the cell site to its homing central office, including the central office itself, and 3) loss of power to the cell sites. Of these, number 3 had the largest impact because in the hardest hit Northeastern states, power outages of 1-3 weeks were not uncommon in many areas, and many carriers only provided backup power systems (such as batteries or self-starting generators) for short periods. Gasoline-powered backup systems were a further issue because gasoline shortages became a serious problem due to power outages at gas stations and throughout the gasoline supply chain. From a national perspective, this is a problem that melds power issues with telecommunications issues. For example, what kind of economical backup power systems could be deployed at tens of thousands of cell sites? How would the backup systems themselves be immune from such natural disasters, and how much would these backups increase the probabilistic "uptime" of cellular service against a model of disasters? This model would have to weigh cost vs. increased network availability.

T for Time

Last but not least, the dimension of time needs to be factored into complex engineered networks in several ways.

First, metrics along the time dimension should be part of the design goals in complex engineered networks. For example, delay, deadline, or cascades of deadlines (e.g., in interactive services) need to become part of the objective in the control and optimization of complex engineered networks, in addition to throughput, distortion, and energy. This tends to introduce both the difficulty of modeling queues and the tradeoff between time-related metrics and non-time-related metrics. As mentioned in earlier sections, many fine-granular questions in queueing theory remain open while approximations sometimes lose predictive power. Much more science is required before we can engineer glitch-free, real-time communication systems, essential in communication networks, sensor networks and smart grids.

Second, the transient behavior of distributed control in a complex engineered network demands much more attention than it has received so far. From the rate of convergence to the evolution of network states, there is little mature machinery to analyze or synthesize for transient behavior. Yet in real networks, almost all the time is spent on transient stages while equilibrium remains a conceptual ideal. We need to shift some of the focus of building tractable yet still realistic methodology from the convergence to an equilibrium point to the invariant states during transient stages.

Phenomena Leading to Network Impairments

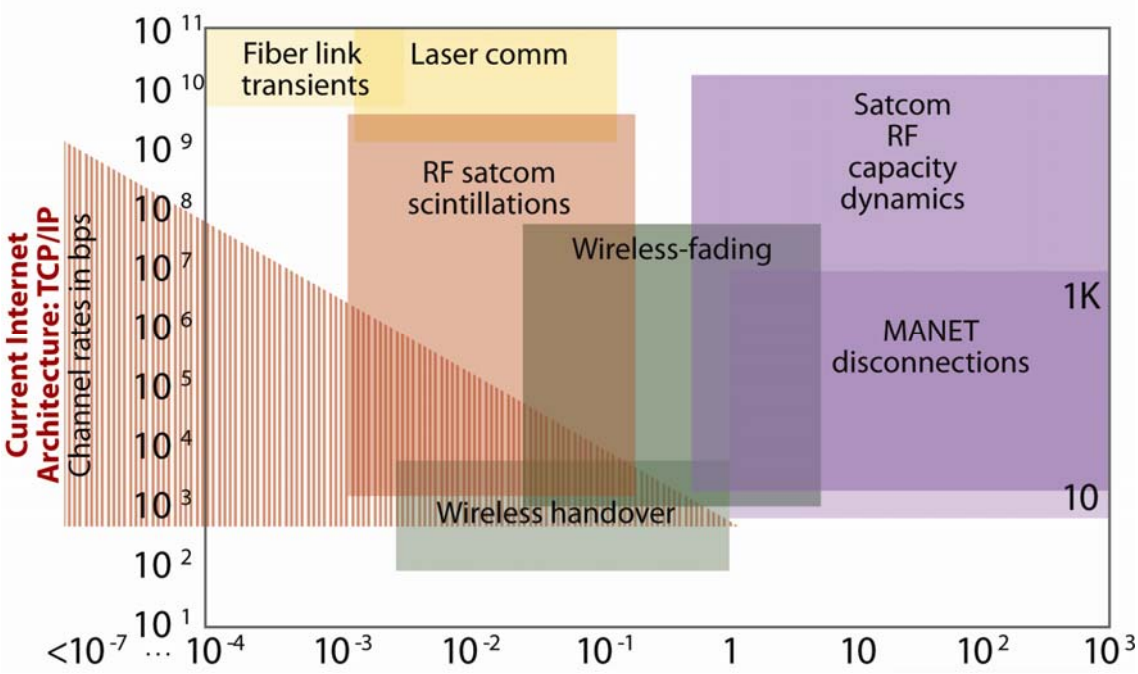

Figure 4: Diverse timescales of network impairment phenomena

Third, while supporting real time application is hard enough, enabling real time management of a complex engineered network is even harder. For example, how can network operators carry out real time (and large-scale) network measurement, learning, and visualization [R+11]? How can fast sensing be carried out with limited sensed information for real time action and resource allocation? Part of the challenge resides in the computational challenges for real-time processing, while other parts are due to the complexity of control plane. As discussed under "Limits," Bode-type fundamental limits that explicitly builds on delay properties are sorely needed to guide real-time control in cyber-physical networks, where excessive delay can be as bad as failures and reliable delivery of deadline is a paramount concern.

Nugget Case Study: Interactive Applications in the Cloud

Because of the flexibility and availability that the cloud offers, more than 30 percent of enterprises worldwide use at least one cloud-based solution. Cloud revenue is expected to grow 500 percent from 2010 to 2020 as cloud applications and companies multiply and expand. Despite the cloud opening so many possibilities, it is not always able to deliver on performance demands, sometimes leading to sub-par end user experiences. For example, research from Google Chrome found that 20 milliseconds of network latency could result in a 15 percent decrease in page load time. Other studies from Amazon and Google found that a half-second delay causes a 20 percent drop in traffic on Google, and a one tenth of a second delay can lower Amazon's sales by 1 percent. Clearly, latency is not only a nuisance, but also a serious problem for enterprises that house their applications in the cloud.

There are three important reasons for this latency. Most networks are broken down into many subsystems that are each owned, operated and managed by different entities. Therefore, enterprises often do not have insight into the performance of their network, let alone the ability to optimize its performance or reduce latency.

Second, compounding the effects of distributed computing, virtualization adds another layer of complexity to cloud latency. Once a simple storage warehouse for rack-mounted servers, today's data centers are a complex web of hypervisors running dozens upon dozens of virtual machines. Within this forest of virtualized data centers, servers often incur packet delays before data even leaves the rack itself.

Third, when it comes to cloud transactions conducted over the Internet, service providers often do not establish SLAs. This is largely because connectivity providers are still working out how they can ensure strong uptime for cloud applications, not to mention what levels to set. Understanding these three problems will go a long way in improving interactive cloud solutions.

Nugget Case Study: Smart Grid Dispatch

In the past few decades, many countries have made an increasing push to integrate renewable energy generators into the grid. These efforts have ranged from large-scale, massive offshore wind farms with capacities of hundreds of MW to local photovoltaic module arrays of hundreds of kW. With these types of energy come problems with transient behavior, because their output levels vary significantly over time. As a result, predictions have to be made hours and even days in advance as to their available power outputs, so that thermal, nuclear, or coal plants can be scheduled to dispatch as much power as necessary to compensate for differences between supply and demand.

This dispatch problem is an optimization solved by the independent system operators (ISOs), who attempt to minimize generation costs subject to flow constraints. It is imperative for operators to model the transients of the grid as accurately as possible. This not only applies to renewable energy forecasting, but also to load demands. The more useful these models are, the less the chance to over- or under-predict the necessary dispatch. In particular, the latter implies that supply is less than demand, meaning certain areas will have their loads curtailed. The areas which are curtailed could potentially be determined by customer contracts with the ISOs. Many grid operators, like PJM and NYISO, have recently begun considering the transient dispatch problem with renewables.

Part III
From Fundamental Research to Societal Impact

Significant investment by the government and industry is continuously made to the modeling, analysis, design, and deployment of complex engineered networks. A range of approaches can be exercised to further facilitate the translation and amplification of fundamental research to societal impact through public-private partnership.

- Encourage failure. Negative results from a grant might be viewed positively by the program manager if they reveal new insights about how not to understand complex engineered networks.
- Encourage falsification of theories. Many theories are developed to be self-consistent while making a long list of assumptions. The resulting predictive power needs to be verified.
- Encourage theoreticians and experimentalists to team up. Mathematical rigor and practical relevance do not have to trade-off against each other. The complete cycle of modeling, design, implementation, deployment, data-collection, and back to sharpened models requires a team of experts with complementary backgrounds, including both the mathematical side of engineering and the systems side.
- Encourage industry and academia to work together. There has been a series of high-profile successes crossing the boundary of industry and academia in the history of complex engineered networks, e.g., cellular networks and the Internet. Special incentives can be provided to encourage more of such collaboration.
- Bridge the theory-practice divide, by facilitating conversations about differences in academia vs. industry culture, including differences in metrics and assumptions, in implementation and management complexities, and in prioritization of problem formulations.
- Leverage shared infrastructure. Going from computer simulation to at-scale experiments takes a tremendous amount of capital, time, and human resource. Shared infrastructure, as well as software tool development, financed by federal funding agencies will go a long way in bridging that gap.
- Encourage the creation of realistic, scalable, and rapid prototyping experimental facilities, and the completion of a research cycle that includes modeling, theory, implementation, data collection, and back to model refinement.

- Create common databases. A repository for data sharing needs to be created in each of the key industry sectors of complex engineered networks, statistically similar to the empirical data but without revealing any privacy-sensitive or company-proprietary information.

- Participate in national policy debates. This is particularly important as complex engineered networks take a center stage on topics such as national energy policies, wireless spectrum policies, and freedom of information access as a foreign policy cornerstone.

- Encourage a non-traditional path to societal impact, including contributions to standards and public policies. Prepare faculty members for the differences in culture, for the time commitment, and for the effective paths of communication in participating in such activities.

- Finance more patent applications. A more relaxed policy or incentive to encourage patent filing based on fundamental research results, together with a more proactive pursuit of patent application and licensing potential, needs to be explored in conjunction with university technology licensing offices.

- Revisit engineering education and workforce development. Develop multi-disciplinary curricula and training mechanisms to ensure adequate workforce. Include more material, possibly case studies, on the process of turning research into engineering artifacts and further into positive impacts on economic productivity in the society. For example, Massive Open Online Courses (MOOCs), delivered through effectively a complex engineered network of learning, offer an interesting potential as well as significant challenges in scaling up effective teaching.

Appendix I
Grand Challenges in Wireless Networks

The Wireless Breakout Session identified three grand challenges and associated opportunities, as well as recommendations for methodologies and technology transfer thought to be critical for overcoming these challenges. The grand challenges and opportunities are as follows:

I. Scaling to Serve 1000x Demand in 10 years. The demand for wireless bandwidth will continue to accelerate throughout the foreseeable future, stimulated by increasing transport of video and high-definition images. Mobile data usage has roughly doubled each year since 2007, and this exponential growth is projected to continue for the foreseeable future. This growth is driven by increasingly powerful user devices as well as bandwidth intensive applications, such as mobile video, mobile cloud, and apps that require continual interaction with remote servers (e.g., Siri). For example, 30 seconds of 1080p video taken from an iPhone 5 will produce a 75 MB data file, which dwarfs the files produced by the iPhone 3G camera (still images only, roughly 700 kB). New research is needed to define an architectural and analytical foundation for supporting dense, heterogeneous network deployment (femto/pico/small cells, relays, message ferrying), including management of interference problems. Algorithms and protocols must be developed to enable fast time-scale access to leverage spectrum gaps opportunistically. New data-driven mathematical models for mobility, connectivity and usage are required, to facilitate quantitative evaluation of new architectures and algorithms.

II. Automation and Self-Optimization from User's Perspective. Research is required to advance toward a truly automated, situation-aware wireless client: optimal, dynamic selection among multiple radio interfaces, where the selection matches delay, throughput and availability characteristics to user and application needs. To fulfill this goal, new methods are needed to ensure both local optimality and global optimality (social welfare) in devices' decision making. At the same time, the resulting architecture cannot significantly add complexity to network management, which would be unattractive to network operators.

III. Network Security and Resilience. The proliferation of smart phones with known security vulnerabilities further expands the growing list of known threats to wireless networks, such as jamming and other denial of service mechanisms, man-in-the-middle attacks, and compromise of base stations. Threats to the confidentiality and integrity of user data are similarly multiplied. New research is needed to secure the network against attack from rogue actors (internal and external), including rapid diagnosis and mitigation within an increasingly distributed architecture. The evolution of wireless architectures toward small-cell designs presents both challenges and opportunities along these lines.

The current art's inability to address these challenges results, in part, from the need for new methodologies to guide the research. The definitions of better metrics and data analytics methodologies for performance assessment are critical for understanding sub-optimal behaviors in wireless network architectures. Similarly, research efforts must devise and implement instrumentation to provide measurement data for assessment and optimization, including existing networks (for example, the use of crowd sourcing and the "mobile cloud" itself to collect measurement in a distributed fashion). Methods for modeling, coordinating, predicting, and evaluating distributed control loops are critical, but are lacking in the current art. Development of more complete means for modeling wireless networks, particularly with respect to network uncertainties and time variations, is also needed.

The transfer of results from wireless network research to practical applications can be greatly facilitated through increased emphasis in three areas. First, research programs should promote and enable early, frequent testing in realistic environments. Second, the research should incorporate input from stakeholders concerning datasets, CONOPS, and other major constraints on ultimate usage (cost, platform, backward compatibility, powering,…). Third, investment should emphasize open platforms, testbeds and spectra that the community can readily utilize.

Appendix II
Grand Challenges in the Internet

The Internet Breakout Session categorized the questions and challenges for the Internet into the following topic areas:

I. Definitions:

Perhaps the biggest challenge to this session is to define the term **Internet**. One of the key takeaways of the session is that this term means different things to different people. However, despite that hurdle, we defined a few broad categories below. Note this does not cover all the possible interpretations of the Internet:

Public Internet. The original incarnation of the "Internet" in the USA has evolved into a collection of independent, commercial carriers (called Internet Service Providers - ISPs) who interconnect via pair-wise peering agreements and protocols (such as BGP). This collection of packet networks allows open IP addresses and reaches all consumer and some business end users (often called eyeballs).

Virtual Private Networks (VPNs). This category supports independent business applications and enables private IP addresses and various forms of privacy and security beyond the public Internet. However, note that often for economic purposes, packets of the public Internet and VPNs mix on the same links of ISP networks. Various protocols, such as MPLS or Ethernet tunneling, can be used to maintain a virtual separation.

Private Packet Networks. These networks are heavily firewalled (both physically and virtually) from the above networks and usually built over a combination of separate routers/switches and lower-layer infrastructures (such sd high-rate private lines, multiplexing and cross-connect equipment, leased dark fiber, or privately-owned fiber). Many mission-critical networks are of this category.

Research and Education Networks (RENs). These networks are often partially subsidized by government bodies (e.g., state, local, or federal) and have objectives to support educational and governmental advanced research/science projects. Some examples are ESnet (sponsored by the Department of Energy) and the Defense Research and Engineering Network (DREN).

II. Security

Cyber Security is a key priority of many US Federal government agencies today. Some of the challenges identified in the session were: How do we handle distribution of malware? What is the degree of dependence among mission critical systems, such as power grids,

transportation systems, network control planes, and financial systems, on the various manifestations of the Internet?

What should be done to better secure access to the above systems, particularly in the context of Cyber Warfare, which can most often be characterized as a foreign security threat?

III. Privacy

Clearly, the issue of privacy is more involved than the simple and verbal privacy policy statements that many websites publicize, which are mostly legal disclaimers. What should individuals, online services, regulators, and policymakers do beyond the policy statements to ensure privacy is protected in this age of big data? A particularly hot nugget question arises as more user information is tracked (often for purposes of marketing, advertising, geo-location applications): how do we prevent or control data mining by third parties to profile individuals? For example, tracking the locations that an individual visits with his/her cell phone can provide a lot of information about that person's purchasing, work, and entertainment habits.

IV. Reliability, Performance, Prediction

One of the largest challenges for designing an ISP network is how to economically design and operate the networks, yet provide a desired level of reliability and performance. For example, it must mitigate the impact of outages due to equipment and facility events (e.g., equipment component outages, fiber cuts, network maintenance, equipment/software upgrades) vs. external events (e.g., natural disasters, acts of war, cyber attacks, such as Distributed Denial of Service - DDOS). The first class of outages is generally more predictable (and hence easier to model) than the latter class. How an ISP handles these two classes of outages differs based on the type of Internet. For example, Private Internets, especially those that may carry government traffic (such as the Defense Information Systems Agency (DISA)) perhaps may emphasize the latter class outages more than commercial ISPs.

Then, extending beyond the question of how to design an individual ISP network, because of the independent nature of ISPs (yet cooperative responsibility via peering relationships), there is often no overall methodology to control instability as traffic jumps among peering points. In fact, it can sometimes be to the advantage of one ISP to shift traffic to another peering point of its neighboring ISP. The latter issue has not generally been addressed and perhaps could profit from some form of societal benefit analysis.

V. Public Policy

The commercial ISPs that comprise the public Internet are largely unregulated (i.e., from governmental oversight) on how they manage policy. Policy can include issues such as pricing, bandwidth policing, definitions and implementation of Class of Service (CoS) and associated Quality of Service (QoS). A potential role for government and the research community is to

assess how the current decoupled structure of the Internet compares to a structure where the Internet is optimized as a whole, i.e., from an overall societal viewpoint. Various issues have arisen in this regard, such Net Neutrality and the FCC's National Broadband Plan. A particularly challenging objective has been how to provide high capacity broadband Internet service in rural and economically depressed areas. Historically, these areas have proven financially challenging to commercial ISPs.

VI. Economics and Pricing

The independent, commercial architecture into which the public Internet has evolved requires consideration of the cost constraints and financial health of the major ISPs who constitute the Internet. Thus, while the critical questions and challenges discussed above need to be explored, they cannot be addressed in isolation and disjoint from economic models. The Internet Breakout session noted a very key observation that the cost and revenue of the Internet is dominated by the last mile or "access or edge network" to get to the end users (eyeballs). Generally, the design and operation of the core (inter-city) network follows changes to the edge; put another way, the edge is where the "action" is.

An important aspect of economics is to better understand the incentives that spur innovation. A particularly hard question is how do we gracefully evolve to such new architectures, given the existence of a heavily and privately-invested embedded infrastructure.

Finally, how should the above cost consideration relate to pricing models in the ecosystem consisting of ISPs, content and app providers, and consumers? This also relates back to the policy debate discussed above.

VII. Methodology

What methodologies can we use to better understand and study these questions? For example, what similarities are there between the types of problems and economic constraints of power grids and transportation networks vs. those of the Internet? With the answers to this question, a follow-up question arises: are there methodologies with synergies for the research and engineering community to exploit?

One related problem identified in this session is that the data inherent to an ISP's services and network are often kept proprietary to that ISP. How can ISPs better release and expose usage data, internal cost models, traffic considerations, architectures, and implementation constructions to the academic and government research community so that they can better participate in its evolution and optimization?

Beyond traffic characterization, design, and architecture of ISP networks, a major issue is the evolution of their control plane. For example, to make networks more cost effective, more reliable, and consume less power, future networks will likely need to be more agile in adapting

to exogenous traffic and dynamic changes in network states. This will put tremendous stress on the control plane. Not all network states can be sensed and used for control in the complex network of the future. A challenge is to elevate Network Management and Control to more of a science so that the architecture can be optimized. For example, what should be the role of software defined radios, switches, and networks and how does this trade off with security and proprietary concerns?

Appendix III
Grand Challenges in Cyber Physical Networks

An engineered network is one that delivers services in a dependable way, whether it is dealing with bits or knowledge or decisions. What distinguishes cyber-physical networks is that they have to provide services to components of a system constrained by the laws of physics, while considering the requirement for storage, computing, communication, and actuation of the underlying substrate.

In a cyber-physical network, the decision paradigm may range from a single decision maker to a distributed decision process. The spatial extent ranges from global to the smallest level within a device. The response time needed also varies significantly. For example, in energy networks, 4-8 ms may be needed for distance protection, differential protection, and over current protection, while 60-100 ms might be necessary for load shedding. In terms of bandwidth requirement, it ranges from high (hierarchical data acquisition and transfer, or strategic power infrastructure defense coordination), to medium (fault event recorder, control devices), to low (generator control, bus configuration, load shedding).

Cyber-physical networks have become an important part of the society. For example, each year $18-20 billion is lost by under-optimizing energy networks, and an estimated $49 billion lost in annual outages can be saved. Each year the nation is investing significantly in opportunities in cyber-physical networks, including smart grids (models, sensors, measurement, methods, and management, etc.), sensor networks (with applications to defense, industry, instrumentation, environment, etc.), and broad applications such as telematics, healthcare, and emergency communications.

I. Theory: A significant challenge is the need for proper metrics and mathematical foundations for understanding, synthesizing, and evaluating architectures in a systematic way, including issues such as robustness, resilience, security, reliability, fragility, and volatility, all of which require formalisms and validation. Similarly, we need a unified theoretical foundation for different network structures, e.g., interdependence of different components of power grids, cyber security, and social network behavior. The desired outcome would be a unified framework and language for heterogeneous architectures, models, and components, with a positive spiral across metrics, theoretical bounds, and empirical data.

II. Resilience: Assuming that network failure will be the norm, we need self-optimization and self-healing capabilities, and a deep understanding of the propagation of risks and uncertainty across subsystems of various scales. These networks need the ability to make decisions at the

speed of a machine and to do so automatically. We need the architecture that can help isolate damage and fortify the network, reduce attacker pool and reduce human error.

III. Tools: We also need scalable tools and methods for understanding, designing, and conducting tradeoffs between various characteristics of the network, including suitability for use, economics consideration, and performance. When it comes to specific applications, we need to create systems and methods for non-intrusive instrumentation and testing, using examples from other fields such as biologically inspired architectures. The desired outcome is a collection of techniques for model abstraction which drives the acceleration of system design.

In facing the above challenges, communities need to be formed, with partnership fostered, common database built, and testbeds shared for networks such as self-healing smart grids and patient-centered healthcare delivery systems.

References

[ACLY00] R. Ahlswede, N. Cai, S.-Y.R. Li, and R. W. Yeung, "Network information flow," IEEE Transactions on Information Theory, 46(7), pp. 1204-1216, July 2000.

[A01] M. Amin, "Toward self-healing energy infrastructure systems," IEEE Computer Applications in Power, 14(1), pp. 20 – 28, January 2001.

[ABKM01] D. G. Anderson, H. Balakrishnan, M. F. Kaashoek, and R. Morris, "Resilient overlay networks," SOSP 2001.

[AM08] K. J. Astrom and R. M. Murray, Feedback Systems: An Introduction for Scientists and Engineers, Princeton University Press, 2008.

[B+10] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," Nature, 464, pp. 1025 – 1028, 2010.

[BAM09] T. Benson, A. Akella, D. Maltz, "Unraveling the complexity of network management," USENIX NSDI, April 2009.

[BT97] D. Bertsekas and J. Tsitsiklis, Parallel and Distributed Computation: Numerical Methods, Athena Scientific, 1997.

[BV04] S. Boyd and L. Vandenberghe, Convex Optimization, Cambridge University Press 2004.

[C13] G. K. Cambron, Global Networks: Their Design, Engineering, and Operations, Wiley 2013.

[CRT06] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Transactions on Information Theory, 52(2), pp. 489-509, February 2006.

[CK74] V. Cerf and R. E. Kahn, "A protocol for packet network intercommunication," IEEE Transactions on Communications, 22(5), April 1974.

[CLCD07] M. Chiang, S. H. Low, A. R. Calderbank, and J. C. Doyle, "Layering as optimization decomposition," Proceedings of the IEEE, 95(1), pp. 255-312, January 2007.

[C53] C. Clos, "A study of nonblocking switching networks," Bell Systems Technical Journal, 32(2), pp. 406 – 424, 1953.

[CT06] T. M. Cover and J. Thomas, Elements of Information Theory, 2nd Ed., Wiley, 2006.

[D06] D. Donoho, "Compressed sensing," IEEE Transactions on Information Theory, 52(4), pp. 1289-1306, April 2006.

[DC00] J. C. Doyle and J. M. Carlson, "Power laws, highly optimized tolerance and generalized source coding," Physical Review Letters, 84(24), pp. 5656-5659, 2000.

[FS11] K. R. Fall and W. R. Stevens, TCP/IP Illustrated, Volume 1: The Protocols, 2nd Ed., Addison Wesley, 2011.

[FM10] T. Flick and J. Morehouse, Securing the Smart Grid: Next Generation Power Grid Security, Syngress, 2010.

[FERC08] Federal Energy Regulatory Commission, Assessment of Demand Response and Advanced Metering, 2008.

[FJ93] S. Floyd and V. Jacobson, "Random early detection gateways for congestion avoidance," IEEE/ACM Transactions on Networking, 1(4), pp. 397 – 413, August 1993.

[H99] C. Huitema, Routing in the Internet, 2nd Ed., Prentice Hall, 1999.

[J+12] V. Jacobson, et al., "Networking named content," Communications of ACM, 55(1), pp. 117-124, January 2012.

[KMT98] F. P. Kelly, A. Maulloo, and D. Tan, "Rate control for communication networks: Shadow prices, proportional fairness and stability," Journal of Operations Research Society, 49(3), pp. 237 – 252, March 1998.

[K75] L. Kleinrock, Queueing Systems, Volume 1: Theory, Wiley 1975.

[KM03] R. Koetter and M. Medard, "An algebraic approach to network coding," IEEE/ACM Transactions on Networking, 11(5), pp. 782—795, October 2003.

[LTWW94] W. E. Leland, M. S. Taqqu, W. Willinger, D. V. Wilson, "On the self-similar nature of Ethernet traffic," IEEE/ACM Transactions on Networking, 2(1), pp. 1 – 15, February 1994.

[LDP02] S. H. Low, J. C. Doyle, and F. Paganini, "Internet congestion control," IEEE Control Systems Magazine, 21(1), pp. 28 – 43, February 2002.

[LMR04] J. C. S. Lui, V. Misra, and D. Rubenstein, "On the robustness of soft state protocols," IEEE ICNP, 2004.

[M00] S. Morris, "Contagion," Review of Economic Studies, 67, pp. 57 – 78, 2000.

[N10] M. Newman, Networks: An Introduction, Oxford University Press, 2010.

[NRC12] NRC Report by Committee on Enhancing Robustness and Resilience of Future Electrical Transmission and Distribution, Terrorism and the Electric Power Delivery System, 2012.

[P+06] L. Peterson, et al. “GENI design principles,” IEEE Computer, pp. 102-105, September 2006.

[R+11] M. Roughan, W. Willinger, O. Maennel, D. Perouli, and R. Bush, “10 lessons from 10 years of measuring and modeling the Internet’s autonomous systems,” IEEE Journal of Selected Areas in Communications, 29(9), pp. 1 – 12, September 2011.

[SJHC13] S. Sen, C. Joe-Wong, S. Ha, and M. Chiang, “A survey of broadband data pricing: Past proposals, current plans, and future trends,” ACM Computing Surveys, 2013.

[V+95] T. Vicsek, A. Czirok, E. Ben Jacob, I. Cohen, and O. Schochet, “Novel type of phase transitions in a system of self-driven particles,” Physics Review Letters, 75, pp. 1226 – 1229, 1995.

List of Participants

Robert Kahn
Mihai Anitescu (Argonne National Lab)
Tim Ashenfelter (DHS)
Isabel Beichl (NIST)
Bob Bonneau (AFOSR)
John Chapin (DARPA)
Chris Dabrowski (NIST)
Darleen Fisher (NSF)
Keith Gremban (DARPA)
Sandy Landsberg (DOE)
Shuai Lu (Pacific Northwest National Lab)
Bryan Lyles (NSF)
Rabi Madan (ONR)
Thomas Ndousse-Fetter (DOE)
Richard O'Neill (FERC)
Nagi Rao (ORNL)
Guna Seetharaman (AFRL)
Blair Sullivan (ORNL)
Bruce Suter (AFRL)
Richard Vojtech (DHS)
Jack Brassil (HP)
Sudip Charkarabati (Osage partners)
Mike Cook (Comcast)
Robert Doverspike (AT&T Labs)
Adam Drobot (Formerly Telcordia)
Chip Elliot (BBN)
Victor Glass (National Exchange Carrier Asso.)
Tom Markham (Honeywell)
Roger Neal (NYC Media Lab)
Jason Li (IAI)
Jin Li (Microsoft)
Yatin Mundkar (Artiman Ventures)
Chris Ramming (Intel)
David Rosenbluth (Lockheed Martin)
Stan Russo (SES)
Krishan Sabnani (Alcatel Lucent Bell Labs)
John Smee (Qualcomm)
Stu Wagner (ACS)
Massoud Amin (U. Minnesota)
Rob Calderbank (Duke)
Vincent Chan (MIT)
Mung Chiang (Princeton)
John Doyle (Caltech)
Robert Fry (JHU Applied Physics Lab)
Mario Gerla (UCLA)
Ali Jadbabaie (Penn)
George Kesidis (Penn State U)
Edward Knightly (Rice)
PR Kumar (Texas A&M)
HT Kung (Harvard)
Muriel Medard (MIT)
Prateek Mittal (UC Berkeley)
Beneditto Piccoli (Rutgers)
Dipankar Raychaudhuri (Rutgers)
Keith Ross (NYU Poly)
Henning Schulzerinne (Columbia)
Ness Shroff (Ohio State)
Junshan Zhang (Arizona State U)

Charge Statement

Charge to Participants: The participants are charged to address the following themes associated with Complex Engineered Networks, including assessing the state-of-the-art research, fundamental challenges, and potential research directions:

- **Understanding of complex networked system behavior**: Complex networked systems may be modeled through ODEs or PDE, optimization/control problems, game/learning/economic models, graph models, etc. The presence of computer networks and codes to monitor and control the physical network adds additional levels of complexity to these models. What are the classes of models and methods that capture the main features of these networks? What are the underlying model scales and discrete-algebraic trade-offs?

- **Robustness and resiliency**: The sheer size and complexity of these networks makes them vulnerable to natural, incidental and intentional failures, some of which could be very complex to quantify and analyze. What are the failure models and analysis methods? What the underlying fundamental robustness and resiliency challenges and solutions? How can we model and design for extremely rare but disastrous events?

- **Bridging the theory-practice gaps**: This workshop will explore the intersection between testing/experimentation and analytical approaches to understanding complex systems. How do we use experiments to inform analysis and how do we use analysis to inform experimental design such that we can obtain deep understanding of complex systems with tractable time and resources? What standard assumptions leading to tractability of models have been hindering the predictive power and application scope of the resulting theory? How can we enhance the feedback loop from at-scale experimental data to theory and models?

- **Transferring fundamental research to societal impacts:** How can federal agencies collaborate amongst themselves and with industry to create pathways for funded research to go from proofs to prototypes? Can the Silicon Valley model be modified to enhance the visible successes from tax dollars to job creation in the networking field?

Agenda
NITRD LSN Workshop on Complex Engineered Networks

September 20-21, 2012, Washington, DC

Wednesday, September 19, 2012

6:00-7:30 PM	Registration

Thursday, September 20, 2012

7:30 AM	Registration
8:30-9:00 AM	Overview by Co-chairs: Mung Chiang and Nagi Rao
9:00-9:30 AM	Sponsoring Agency Perspectives AFOSR: Bob Bonneau ARO: John Lavery NSF: Bryan Lyles DOE: Sandy Landsberg, Thomas Ndousse
9:30-10:30 AM	Rapid Fire Presentations I: Opportunities in Complex Engineered Networks
10:30-11:00 AM	Break
11:00-12:30 PM	Panel I: Areas and Challenges Moderator: Bob Bonneau

12:30-1:30 PM	Working Lunch
1:30-3:00 PM	Breakout Session I Internet Wireless Networks Cyber-Physical Networks
3:00-3:30 PM	Break
3:30-4:30 PM	Rapid Fire Presentation II: Challenges in Complex Engineered Networks
4:30-6:00 PM	Panel II: Foundations Moderators: John Lavery and Sandy Landsberg
7:00-9:00 PM	Working Dinner with Keynote Speaker Robert Kahn

Friday, September 21

7:30 AM	Registration
8:30-9:00 AM	Summary and Overview by Co-chairs
9:00-10:30 AM	Panel III: Industry, Academic and University Collaborations Moderators: Bryan Lyles and Thomas Ndousse
10:30-11:00 AM	Break
11:00-12:00 PM	Rapid Fire Presentation III: Best ways to Research Impact

12:00-1:00 PM	Working Lunch
1:00-2:30 PM	Breakout Session II
	Internet
	Wireless Networks
	Cyber-Physical Networks
2:30-3:00 PM	Break
3:00-4:30 PM	Breakout Session Reports
	Internet
	Wireless Networks
	Cyber-Physical Networks
4:30-5:30 PM	Concluding Discussions