“Shock and Awe:" *Network Science* as an Interdisciplinary Networking Course

Walter Willinger
AT&T Labs-Research
Florham Park, NJ 07932
walter@research.att.com

ABSTRACT

In many respects, the recent popularization of Network Science as "the next big thing" and its impact on much of Computer Science (CS) has many similarities with "shock and awe" – a term used by the Bush administration for its massive hi-tech air strikes at the beginning of the the Iraq war. Telltale signs that CS is currently experiencing some of the symptoms that are typical of "victims" of this "shock and awe" treatment are (i) an uncritical and almost submissive embrace of Network Science concepts, (ii) a disturbing denial of being firmly rooted in the engineering sciences, and (iii) an acute fear of having lost its identity. In this purposefully provocative but completely tenable article, I argue that it's time for CS to turn the tables and show Network Science its proper place.

1. INTRODUCTION

"Shock and awe" was the term used by the Bush administration for its massive hi-tech air strikes at the beginning of the the Iraq war. As a military strategy, it is discussed in detail in [12] where the authors describe it as being *"aimed at influencing the will, perception, and understanding of an adversary rather than simply destroying military capability."* They go on and explain that *"rapid dominance will impose this overwhelming level of Shock and Awe against an adversary on an immediate or sufficiently timely basis to paralyze its will to carry on ... [to] seize control of the environment and paralyze or so overload an adversary's perceptions and understanding of events that the enemy would be incapable of resistance at the tactical and strategic levels."*

The recent popularization of Network Science as "the New Science" and its impact on much of CS has many similarities with this "shock and awe" strategy. Fueled in part by the publication of a National Research Council report in 2005 [8], Network Science has become a new and rapidly evolving discipline that has gained great visibility in the technical and popular literature within the last few years. In a recent article, L.-A. Barabasi [6], one of its founders, summarized the overall impact of Network Science as follows. *"For decades, we tacitly assumed that the components of such complex systems as the cell, the society, or the Internet are randomly wired together. In the past decade, an avalanche of research has shown that many real networks, independent of their age, function, and scope, converge to similar architectures, a universality that allowed researchers from different disciplines to embrace network theory as a common paradigm. The decade-old discovery of scale-free networks was one of those events that had helped catalyze the emergence of network science, a new research field with its distinct set of challenges and accomplishments.* What is certainly undeniable is that the number of published papers in this field has increased at staggering rates [2], workshops and conferences on this topic have mushroomed, courses on different aspects related to this field are offered by various organizations (e.g., [9, 10]), course materials at different levels are readily available on an ever-growing number of websites (see, for example, [5] under "resources"), and new textbooks keep appearing in rapid succession. Some of the more popular books in this area include, among others,

- L.-A. Barabási. *Linked: How Everything Is Connected to Everything Else and What it Means for Business, Science, and Everyday Life*, Perseus Publishing, Cambridge, MA 2002.
- D.J. Watts. *Six Degrees: The Science of a Connected Age*, Norton, New York, 2003.
- S.N. Dorogovtsev and J.F.F. Mendes. *Evolution of Networks: From Biological Nets to the Internet and WWW*, Oxford University Press, Oxford, 2003.
- R. Pastor-Satorras and A. Vespignani. *Evolution and Structure of the Internet: A Statistical Physics Approach*, Cambridge University Press, Cambridge, 2004.
- M.E.J. Newman, A.-L, Barabasi, and D.J. Watts. *The Structure and Dynamics of Networks*, Princeton University Press, Princeton, NJ, 2006.
- A. Barrat, M. Barthelemy, and A. Vespignani. *Dynamical Processes on Complex Networks*, Cambridge University Press, Cambridge, 2008.
- T.G. Lewis. *Network Science: Theory and Applications*, Wiley, 2009.
- M. E. J. Newman. *Networks: An Introduction*, Oxford University Press, March 2010.

Network Science has achieved this "rapid dominance" despite the fact that there still exists a considerable divergence of views regarding the precise definition and overall scope of this new discipline [7]. While the NCR report [8] attempted to provide a working definition (i.e., *"Network science consists of the study of network representations of physical, biological, and social phenomena leading to predictive models of these phenomena"*), most views of the field to date probably agree on three basic features that, to some degree, characterizes this "new science:" the work tends to be data-driven, analytic in nature, and based on the premise that network properties that are universal across very diverse areas of applications exist and await discovery. These views have been shaped by the recent availability of enormous amounts of network-related data from all areas of science, by a largely physics-inspired and -dominated approach to dealing with complex large-scale networks, and a very traditional attitude towards the roles of data analysis, modeling, and model validation in an age of unprecedented access to unprecedented amounts of measurements.

2. NETWORK SCIENCE AND CS

A main reason why CS has been a prime target for Network Science's rapid dominance strategy is that for the past 10 years, the Internet has provided an almost ideal playing ground for Network Science ideas and hence has been a prime application area for many of the concepts advocated by this new kind of science. In fact, the Internet has featured prominently in the early and highly publicized success stories attributed to Network Science [1], partly because of the many types of "networks" or connectivity structures that result from its designed nature (e.g., router-level topology, autonomous system-level topology, the Web graph, Peer-to-Peer networks, Online Social Networks, etc.), partly due to an abundance of readily available high-volume datasets, but mainly because of its importance for our daily lives and its ubiquitous use by billions of people.

This critical and growing dependence of a large fraction of the human population, of almost all organizations, and practically every government on the Internet has attracted increasing attention from different interest groups worldwide and has led to significant funding efforts by public and private institutions alike to improve the overall understanding of this critical infrastructure. Network Science has been quick in recognizing the ensuing opportunities to influence public policy and shape a research agenda in an area that has traditionally been part of the engineering sciences. In contrast, the CS community has been slow to see the writing on the wall and has shown all the signs of a victim of a successful "shock and awe" attack as explained in [12]. For example, only very recently have we seen concentrated efforts for establishing research centers within engineering schools dedicated to performing foundational research on the Internet and other types of communications networks. Another relative recent development within the CS community has been a sudden increase in Network Science course offerings as part of the traditional CS curriculum for undergraduate or graduate students; examples of such recent course offerings include

- CMU: *www.ece.cmu.edu/courses/18799H*
- Columbia University: *www.ee.columbia.edu/ cylin/*
- Georgia Tech: *www.cc.gatech.edu/ dovrolis*
- MIT: *stellar.mit.edu/S/course/6/sp11/*
- UNM/SFI: *www.santafe.edu/media/cscourses/*

However, a cursory overview of these course offerings confirms that most CS departments have largely succumbed to the "shock and awe" strategy applied by Network Science, have given in to the hype or buzz that surrounds this "new science", and have basically been in denial about the engineering roots that attracts so many of the students to CS in the first place. This attitude should be disturbing, not only for networking researchers but also from the perspective of CS educators.

3. CS AND NETWORK SCIENCE

To explain why this attitude should be a cause of concern and why the recent uncritical embrace of Network Science by the CS community sets a bad example, it is necessary to look beyond the existing hype and buzz and dig deeper. While the existing literature is clear about the importance of the Internet application domain for Network Science, there is no denying that the Internet has also emerged as a textbook example for illustrating how and why Network Science has become a classic lesson in how errors of various forms occur and can add up to produce results and claims that create excitement among non-experts but quickly collapse when scrutinized or examined by domain experts. While there are no textbooks that document these failures of applying Network Science thinking to the Internet, there are more and more research papers in the published literature that detail the various mis-steps and show why findings that look at first glance impressive and conclusive to a science-minded reader turn out to be simply wrong or completely meaningless when examined closely by domain experts. A partial list of these papers include

- L. Li, D. Alderson, J.C. Doyle and W. Willinger. A first principles approach to understanding the Internet's router-level topology, in: *Proc. ACM SIGCOMM'04, ACM Computer Communication Review* 34(4), 2004.
- W. Willinger, D. Alderson, and L. Li. A pragmatic approach to dealing with high-variability in network measurements, in: *Proc. ACM SIGCOMM Conference on Internet Measurement IMC'04*, 2004.
- J. C. Doyle, D. L. Alderson, L. Li, S. Low, M. Roughan, S. Shalunov, R. Tanaka, and W. Willinger. The "robust yet fragile" nature of the Internet, in: *PNAS* 102(41), 2005.

- L. Li, D. L. Alderson, J. C. Doyle, and W. Willinger. Towards a Theory of Scale-Free Graphs: Definition, Properties, and Implications, in: *Internet Mathematics* 2(4), 2006.
- D. Alderson. Catching the "Network Science" Bug: Insight and Opportunity for the Operations Researcher, in: *Operations Research* 56, 2008.
- W. Willinger, D. Alderson, and J.C. Doyle. Mathematics and the Internet: A Source of Enormous Confusion and Great Potential, in: *Notices of the AMS* 56(5), 2009.
- D. L. Alderson and J.C. Doyle. Contrasting views of complexity and their implications for network-centric infrastructures, in: *IEEE Transactions on Systems, Man, and Cybernetics-Part A* 40(4), 2010.
- B. Krishnamurthy and W. Willinger. What are our standards for validation of measurement-based networking research? in: *Computer Communications* 34, 2011.
- 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, (under submission), 2011.

While this situation has naturally become a source of great confusion within the larger science community, for CS researchers the main conclusion is neither controversial nor should it come as a big surprise: in its present form, Network science is largely incapable of dealing with designed systems (e.g., the Internet or other highly engineered networks) in a way that advances our understanding of these systems to the point where we can predict their behavior and build them to meet possibly different or new needs. The Internet example also demonstrates the dire need to develop an intellectually stronger Network Science that can pass the more demanding and scientifically more challenging validation criteria required by a more engineering-oriented and less physics-inspired research community.

4. BACK TO THE CS ROOTS

By carefully tracing and documenting the main sources of errors regarding the application of the current Network Science approach to the Internet, the above-mentioned papers show that many of the most popular Network Science concepts are severely lacking in rigor. The main problems include (i) a dismal attitude towards data hygiene, (ii) a largely ignored mismatch between the rigor of statistical data analysis and the quality of the available data, and (iii) an outdated and completely inadequate approach to modeling and model validation when faced with an abundance of data.

Ironically, these problems are all in areas where CS with its roots in the engineering sciences has been traditionally strong. Empirical studies that start with measurements and their analysis and employ creative model building and validation have been at the core of CS's contributions to Internet research. Key to the success of such studies has been the readily available domain knowledge that exists in this field and that Network Science has apparently been successful in ignoring for the sake of focusing on the "big picture." However, it is precisely this domain knowledge that ought to be used to enforce paradigm shifts in areas where Network Science has been caught cutting corners to the detriment of scientific rigor and where it has been successful in convincing large parts of the scientific community that engineers, because of their obsession with details, are largely incapable of "seeing the forest for the trees" and are obstructing the progress of science.

There are three main areas where CS can return to its roots in the engineering sciences and turn the tables on Network Science. They illustrate the key differences in approaches and perspectives that have been responsible for much of the reported "divergence in opinions." An important question they raise is whether Network Science and CS will be able to learn from one another in the future to advance the study of complex networks to the point where it benefits science as a whole.

Measurements: Among scientists, a popular and telling view of the Internet is that since it is a network of computers, and since computers are good at measuring, Internet measurement is easy and data is abundant. However, even within the CS community, a largely ignored fact about Internet-related measurements is that what we can measure in an Internet-like environment it typically not the same as what we really want to measure (or what we think we actually measure). This makes measurement efforts within the larger Internet setting in general non-trivial, and a commonly-used "solution" consists of relying on engineering hacks that typically do not yield the originally desired data and provide instead some closely related but measurable quantities. Using the resulting data at face value (i.e., as if they were the data we originally wanted) and deriving from them results that we can trust can either involve a leap of faith or the use of domain knowledge. The papers cited in Section 3 provide plenty of concrete evidence that while the leap-of-faith method advocated by Network Science often results in wrong results and scientific dead ends, the use-domain-knowledge approach pursued by a CS community that is not in denial of its engineering roots ensures progress and leads to scientific advances. "Measuring the measurer" and "knowing your data" are two scientific activities that lack the allure or appeal of many Network Science concepts but are essential for ensuring scientific rigor.

Data analysis: In addition to the fact that Internet-related measurements typically reflect what we can measure and not what we want to measure, they also have the problem that they tends to be inaccurate, incomplete, or ambiguous. When faced with such data, it is important to remember that there are critical differences between analyzing high-quality and low-quality data sets, and that approaching the latter the same way as the former is not only bad statistics but also bad science and bolsters the popular notion that “there are lies, damned lies, and statistics.” Unfortunately, the work required to analyze such data and arrive at conclusions that we can trust is hardly glamorous or news-worthy, especially when compared to the overall excitement generated by the popular Network Science perspective that emphasizes the enormous volume of the available datasets and their apparent complexity, but is largely agnostic not only to the approximate but also to the often very low-quality nature of the data. Again, ensuring that the “Garbage In, Gospel Out" extension of the phrase “Garbage In, Garbage Out" does not apply to the analysis of Internet-related measurements requires attention to details and "dirty work", a message that should come loud and clear from CS, because it is clearly not part of today’s Network Science teaching.

Modeling: While the Internet application has been a prime victim of the "shock and awe" assault of Network Science, it has fortunately also been a perfect example for demonstrating an alternative approach to the study of networks that highlights the sort of paradigm shifts needed in our quest for an intellectually stronger, mathematically more solid, and scientifically more rigorous "Science of Networks." A key element in this quest is the recognized need to turn network modeling from a largely uninspiring and often flawed exercise in data-fitting into a more challenging but also more rewarding exercise in "reverse-engineering" – the ultimate goal of an engineer to discover and understand the various principles underlying the system of interest. In the process, the focus of Network Science will naturally turn to the networks’ various purposes and functionalities and away from the current almost exclusive emphasis on topology or connectivity. This alternative approach is deeply rooted in the CS community’s engineering orientation and re-iterates the central role played by domain specific knowledge – "details matter!" In this sense, it provides an important balance to the existing, more physics-centered Network Science perspective that seeks the discovery of properties that are universal across a range of diverse networks and do not depend on the particulars of the systems at hand.

5. MOVING BEYOND "SHOCK AND AWE"

There is no reason to believe that because of this engineering-inspired perspective advocated in this article and the increased attention to domain-specific details that it demands, the quest for abstractions of understanding that are common among networks across widely separated domains and that together may capture the essence of broad classes of networks has to be abandoned. In fact, the Internet example and resulting engineering perspective serve as important reminder that while the quest for a unifying theoretical framework that encompasses, builds on, and integrates all these abstractions simultaneously remains well within the reach of a long-term research program in Network Science, the form that these abstractions can take and their relative importance will necessarily differ for networks in different domains or even within the same domain. Clearly, such a fundamental framework would have tremendously broad applicability and appeal, but to get there, things have to change.

To quote from the introduction to a recent special issue of Science [11]: *"In the past 10 years, new ways of gathering, analyzing, storing, and disseminating information have transformed science. Researchers generate more observations, more models, and more automated experimentation than ever before, creating a data-saturated world. The Internet has changed how science is communicated and given non-scientists new opportunities to take part in research. Whole new fields, such as network science, are arising, and science itself is becoming more of a networkŮmore collaborative, more multidisciplinaryŮas researchers recognize that it takes many minds and varied expertise to tackle complex questions about life, land, and the universe ..."* While there is little to nitpick about this view, we argue here that to succeed in this endeavor, Network Science has to learn from past mistakes, broaden its view, and build deeper links with domain experts across the sciences. Otherwise, the same or similar mistakes that we have encountered in the Internet context will be (and are already) repeated in the context of, say, biology, where the implications are potentially much more grave (e.g., people die), and this will reflect poorly on Network Science as a scientific discipline and on science as a whole.

6. REFERENCES

[1] R. Albert and A.-L. Barabasi. Error and attack tolerance of complex networks, *Nature* 406, 2000.

[2] D. Alderson. Catching the "Network Science" Bug: Insight and Opportunity for the Operations Researcher, *Operations Research* 56, 2008.

[3] D. Alderson. podcast, www.scienceofbetter.org/podcast/alderson.html.

[4] D. Alderson. Interview for ACM Ubiquity ubiquity.acm.org/article.cfm?id=1595423

[5] Barabasi’s lab at Northeastern, http://www.barabasilab.com/.

[6] L.-A. Barabasi. Scale-Free Networks: A Decade and Beyond, *Science* 325, 2009.

[7] R. Guerin. What is Network Science? www.seas.upenn.edu/ guerin/presentations/

[8] National Research Council. *Network Science*, The National Academies Press, Washington, D.C., 2005.

[9] New England Complex Systems Institute (NECSI). www.necsi.edu/.

[10] Santa Fe Institute. www.santafe.edu/.

[11] Insights of the Decade, *Science* 330, Dec. 2010.

[12] H. K. Ullman and J. P. Wade. *Shock and Awe: Achieving Rapid Dominance*, National Defense University Press, October 1996.