This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

The Role and Potential Impact of Social, Behavioral and Economic Science Approaches to Networks in the First Half of the 21st Century: Grand Challenges of Substance and Methods*

Bernice A. Pescosolido and Ann McCranie

Indiana University

*White paper prepared in response to the National Science Foundation call (NSF10-069) about grand challenges for future research in the social, behavioral and economic sciences. We thank Scott Long and Jack K. Martin for their comments along the way and the Indiana Consortium for Mental Health Services Research and the College of Arts and Sciences, Indiana University, for support. This paper is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 3.0 Unported License. To view a copy of this license, visit http://creativecommons.org/licenses/by-nc-sa/3.0/ or send a letter to Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

The Role and Potential Impact of Social, Behavioral and Economic Science Approaches to Networks in the First Half of the 21st Century: Grand Challenges of Substance and Methods

Abstract

The last decade has witnessed a path-breaking convergence between natural/physical, computing and SBE sciences in the adoption of a network approach to understanding phenomena from protein structures to global transportation connections. Although social sciences have been involved in the study of networks for nearly a century, recent innovations have come from physics, cognitive, computer and information sciences. While promising, the integration of insights from SBE sciences into the new Network Science is far from complete, not yet incorporating the full range of ideas and approaches. We introduce three grand challenges that stand in the way of Network Science's potential to unlock the fundamental workings of natural, social and artificial systems.

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

The Role and Potential Impact of Social, Behavioral and Economic Science Approaches to Networks in the First Half of the 21st Century: Grand Challenges of Substance and Methods

"From chaos comes complexity. From the movement of molecules within our cells to communication across an entire planet, we are part of networks.…We need more and better data in many disciplines. It is not enough to look at patterns; we need to study how they evolve and change. The magnitude of the challenges we are facing shows how much we still need to learn…" Science *Editorial*, 7/24/ 2009

INTRODUCTION

A series of recent reports from the National Science Foundation (2006), National Institutes of Health (2008), National Research Council (2001) and Institute of Medicine (2000) have highlighted two fundamental directions for future scientific progress – complexity and transdisciplinarity. The first signals the growing awareness that understanding phenomena from diseases to decisions cannot be achieved through a simple lens that examines one level or influence whether biological, social, technological or physical. Rather, a focus on the interconnections among a large number of *interacting* units represents a more realistic and promising way to understand systems and actions where explicit blueprints and self organization are absent (Ostrom 2009). The second suggests that the historical portioning of areas into disciplines can no longer serve as the only, or even best, road to continued research progress.

By working across traditional boundaries, researchers have begun to accelerate the integration of theoretical ideas and research technologies under the set of ideas and approaches that have recently been referred to as Network Science. The rich history of a century of social science research on

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

networks combined with new approaches, algorithms and computing power offers a potentially powerful mechanism to understand the workings of complex systems across broad areas of science, including information and technology; biological systems, health and health care; local and global political and economic processes; and socio-technical systems (Barrat et al. 2010). To date, this approach has produced important scientific findings across substantive areas that presented particularly difficult challenges (e.g., "epidemics" such as SARS, H1N1, HIV, tuberculosis, autism). Many of these advances have coupled development of theory with the advent of new measurement, analytic and visualization tools such as agent-based modeling, microarray analysis, simulation and algorithms for large scale networks.

In sum, the potential for synergistic translation across levels of analysis and disciplinary silos necessary for unpacking complex systems is enhanced by this shared interest in networks. It provides a common starting point that fulfills basic requirements for a transdisciplinary framework. Specifically, integrating the sciences to address fundamental questions rests on the development of approaches that: (1) consider and articulate the full set of contextual levels documented to have impact in past empirical research; (2) offer an underlying mechanism or "engine of action" that connects levels, is dynamic, and allows for a way to narrow down focal research questions; (3) employ a metaphor and analytic language familiar to all sciences to facilitate synergy; and (4) understand the need for and use the full range of proven theoretical, methodological and analytic tools (Pescosolido 2006). Most importantly, unlike other areas of science where investigators from many disciplines focus on the same phenomena (e.g., the successful management of Social-Ecological Systems in Ostrom's research), the applicability of network theory and methods spans a wide range of important scientific and societal problems.

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

This combination of both longstanding and new disciplinary traditions coupled with wide substantive scope is promising and problematic. Four fundamental challenges for the inclusion, participation and leadership of SBE sciences are apparent.

GRAND CHALLENGE #1: ENSURING THE CENTRAL ROLE OF ALL OF SBE SCIENCES IN NETWORK SCIENCE.

While Network Science encompasses all types of research and researchers, the new dominant focus tends toward large scale, complete and mediated networks (e.g., cyber networks of various kinds which are real but not inclusive of all types of connections). The longer and more varied history and range of SBE research on networks has the capacity to offer a theoretical and analytic basis for understanding Network Science commonalities in language and foci *across* topics and approaches. The SBE body of research delineates the central mechanism for the interconnection of both large and small units in complex systems including neural networks, friendship groups, the organization, and global trade structures.

To date, however, the new Network Science has focused primarily on socio-technical structures given the advent of powerful computing and access to large-scale electronic or virtual databases (on citations, internet, social networking). This has produced a great leap in our understanding of networks. However, a number of scientific problems, especially those that deal with more direct human connections, present theoretical challenges in understanding how different types of networks at different levels of analysis are embedded in and influence one another. For example, the meaning of and complicated relationship underlying electronic networks and limits in existing data sources have implications for the understanding of SBE science itself. Studies that use large, existing databases (e.g., the Web of Science) tend to largely exclude large knowledge regions of the SBE sciences. Data sources

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

which truncate understandings of scientific knowledge, dissemination and translation of the SBE sciences mean that our understanding of the landscape of science is incomplete and under-represents the contributions of the SBE sciences.

"Scraping" data from the "socio-technical" world, whether from social media or transportation networks, currently lacks the rigorous evaluation of data quality typical of SBE. Further, these represent only one type of network. A primary focus on virtual ties ignores other networks critical to understanding how ties are forged, maintained, or broken. The multi-faceted nature of ties and their effects on human behavior are often underplayed since large-scale networks often tap one or a few ties (e.g., as complicated as they are, transportation networks are one type of network). Human networks are multi-faceted, often considering structure, function, and content. These require more complicated data structures to understand the operation of social networks at an even richer level than is currently done. To date, little research embeds individual ties in larger contextual analyses.

In sum, the wealth of experience that SBE scientists have in the collection, evaluation and analysis of multi-level data has not been brought to bear on Network Science thus far.

GRAND CHALLENGE #2: ISSUES IN SAMPLING, EGO-CENTRIC NETWORKS AND MISSING DATA.

Many exciting findings in the field of network science in the past decade have employed "whole" or "complete" network analysis, particularly in large-scale networks. This kind of research, however, depends on the clear delineation of study population boundaries and the collection of data about every "actor" in that population and their interactions with all other actors in a population. It also requires minimizing if not eliminating missing data since results in complete network analysis can be seriously affected by missing data. Thus, while human connections are sometimes held within boundaries that can be drawn (e.g., organizational networks), much more often, data relevant to SBE requires the

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

collection of egocentric network data, i.e., actor-based ties networks consisting of actors and their immediate contacts in unbounded contexts. At present the computational and statistical solutions for how to sample network data from the universe of ties that exist outside of the virtual world have not been solved. These will require large-scale efforts in data collection and computing that remain unaddressed.

While many recent advances are applicable to protein networks, computer networks, or transportation networks (where population data can be collected reliably, even if not easily), collecting large-scale networks about many types of human social interaction (friendship and contact, for instance) remains a very difficult and even impossible prospect for SBE researchers who seek to push the boundaries of network science along with their natural science counterparts. Simply put, the standard representative sampling frameworks that SBE has come to rely on to collect data do not work well in a network framework that demands information about *all* actors and their interactions. For example, it is not feasible to interview everyone in even a small-sized town or a large-sized workplace. Even if this limitation could be overcome, we are becoming increasingly aware of the permeability of such boundaries regarding social influence.

Two prevailing solutions have emerged: studies that focus attention on socio-technical relations (described above) and more commonly and historically, studies that draw on smaller, more specific network boundaries with the aim of generalizing to the larger population. Both of these approaches have clear utility in the SBE sciences. But, three areas for methodological advancement would bridge the large-scale network divide and push all types of Network Science forward: the integration of *egocentric* approaches with complete network analysis, development of approaches to *sampling* within networks, and better guides to handling *missing* data in complete network analysis. These issues present vital questions in fields that are beginning to show increased interest in Network Science (e.g., health, military sciences). While there are some promising approaches to data collection (e.g., respondent

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

driven sampling), too often, proposed analytic solutions offer little more than admonitions to minimize missing data and to collect more data. If Network Science is to truly thrive and to take advantage of the advances in large-scale network analysis, practical solutions must be developed for how to handle such limitations and how to integrate smaller portions of larger networks (such as egocentric or sampled portions) into larger network structures.

GRAND CHALLENGE #3: THE DEVELOPMENT OF INFRASTRUCTURE IN NETWORK SCIENCE WITH PARTICIPATION OF AND PRIORITIES FROM THE SBE SCIENCES.

Progress can be hindered by differences in language, traditions, tools and time (Ostrom 2009). Focused efforts on the multidisciplinary development of Network Science, and the important role of the SBE sciences during this early stage are critical. While there are relatively new national and global research initiatives, most of these are focused on socio-technical networks and driven by scientists outside of SBE. The connections often do not directly measure ties between individuals but instead are proxied by systems of transportation or cyberspace. Relatively speaking, existing research infrastructures have not included social networks, in their direct human form, to any large degree.

No center for advanced, transdisciplinary training in Network Science exists; whether or not SBE sciences are included. Pedagogical philosophies and structures that include SBE scientists in key leadership positions are necessary to introduce, integrate and assess transdisciplinary approaches; to create a new generation of scientists who will shape future research agendas; and that will provide a rich archive of network data available to all scientists. Further, social network data have a particular structure and are not currently well-represented in the large data archives available to social science researchers (e.g., the ICPSR). New approaches in Network Science also require computing skills which are not typically acquired, by training or even by interest, by SBE scientists working from a network

This paper was submitted to the National Science Foundation as part of its SBE 2020 planning activity (www.nsf.gov/sbe/sbe_2020/). Its inclusion does not constitute approval of the content by NSF or the US Government. The opinions and views expressed herein are those of the author(s) and do not necessarily reflect those of the NSF or the US Government.

perspective. Linking individual dynamics to system dynamics is a massive task of data collection, compilation and analysis not yet capable of being addressed with current infrastructure limits.

These concerns require infrastructural support for bringing together dedicated researchers who share overlapping scientific concerns, ensuring the participation of SBE scientists. Specifically, it requires developing foundational activities that support the potential for major synergistic efforts; building additional infrastructure support for transdisciplinary teaching, research and diversity activities; the creation of a new social network databank for researchers to easily document and share their databases; and cyberinfrastructure that goes beyond equipment and algorithms to include access to people who have cutting edge, sophisticated statistical software development skills. It is simply not reasonable to assume that well-trained SBE researchers are also to be expected to be software developers. While promising open-source platforms for software development (such as R and Network Workbench) exist, the ability of SBE to add to these platforms is often limited.

REFERENCES

Barrat, A., M. Barthelemy, & A. Vespignani. 2010. *Dynamical Processes in Complex Networks.* New York: Cambridge University Press.

Ostrom, E. 2009. "A general framework for analyzing sustainability of social-ecological systems." *Science* 325: 419-422.

Pescosolido, B.A. 2006. "Of pride and prejudice: The role of sociology and social networks in integrating the health sciences." *Journal of Health and Social Behavior* 47:189-208.