Network Science **1** (1): 1–15, 2013. © Cambridge University Press 2013
doi:10.1017/nws.2013.2

EDITORIAL

What is network science?

ULRIK BRANDES
Department of Computer and Information Science, University of Konstanz, Germany
(e-mail: .b a de @ - a .de)

GARRY ROBINS
Melbourne School of Psychological Sciences, University of Melbourne, Australia
(e-mail: a @ e b.ed .a)

ANN McCRANIE
Department of Sociology, Indiana University Bloomington, USA
(e-mail: a cc a @ d a a.ed)

STANLEY WASSERMAN
Departments of Psychological and Brain Science and Statistics, Indiana University Bloomington, USA
(e-mail: a a @ d a a.ed)

Abstract

This is the beginning of Network Science. The journal has been created because network science is exploding. As is typical for a field in formation, the discussions about its scope, contents, and foundations are intense. On these first eee fi1stEs there a science of networks?

here is or is not a science of networks, and what should be the focus of ence, may appear to be academic questions. We use this prominent place he following three statements:

se questions are more relevant than they seem at first glance.
answers are less obvious than have been suggested.
future of network science is bright.

the editors of a new scientific journal are going to predict that the future cience is bright. Cambridge University Press is betting that it is. We do , however, that the field is still in formation and that its essential shaping place right now. The development and steering of network science into a beneficial direction form our primary motivations for creating this journal suming active roles in its leadership.

k science, the field, permeates a wide range of traditional disciplines, ork Science, the journal, will welcome contributions from all of them. on, we will indeed publish foundational research on theory, principles, y, and mathematics of networks.

While there is much overlap in disciplinary network research, the inherent disciplinary boundaries still tend to create silos of di erent interests, methods, and goals. If network science is to be one science, rather than separate and scattered research communities, or a set of tools that researchers use to analyze networks, the silos need to be dismantled while at the same time recognizing existing disciplinary practices and values.

We write this editorial to help establish common grounds for our journal and its field—grounds which should allow Network Science to excel above and beyond disciplinary boundaries. We envision our commons as much wider than some current interpretations of the term network science. We will therefore try to delineate the uniting elements rather precisely on the next few pages.

Our major statement is that we view network science as the study of the collection, management, analysis, interpretation, and presentation of relational data. But first, a few remarks on our perceptions of the current state of the field.

The claim that "networks are everywhere" has become almost routine. Frequently mentioned examples of "everywhere networks" include the Internet and other infrastructure networks, social, political and economic networks, scientometric and text-representational networks, as well as food webs and molecular-level biological networks. And there is a host of other, less commonly mentioned networks in many more research areas.

Networks and hence the network paradigm have become scientifically relevant across disciplinary boundaries. But many have asked: Is the network paradigm nothing more than an in vogue buzz phrase? Clearly, a science of networks requires of the network paradigms. Political Science is a good example—network science has had a tremendous impact on this field, just in the last five years.

If anything, network science is a revolution a long time in the making. Despite frequent claims by some, network science did not suddenly appear when it was realized in the mid-1990s that networks could be models of complex systems. Such a limited definition of network science is simply inappropriate—it is important to recognize the many scientific antecedents of what we do. Network approaches have developed in many areas over the past two decades (physics, biology, economics, for example) because a relational perspective clearly added relevance to the discipline.

The roots of network science are particularly strong in social psychology, sociology, and anthropology, which has led to another misperception, namely that network science is the application of network analysis in disciplines other than the social and behavioral sciences. Sometimes the phrase social network analysis ("SNA") is used to label everything that is network-related, even when the network aspects of

Fig. 1. The elements of network models.

1. A specification of how the phenomenon (in general, i.e., more generally than this particular instantiation) is abstracted to a network.
2. A specification of how this conceptual network is represented in data (e.g., measured or observed).

As representation is usually defined via an isomorphism, i.e., a one-to-one mapping between structures preserving relations, a phenomenon cannot be represented

According to our framework there are actually two aspects to network theory. On the one hand, network theories can suggest and explicate, for given research domains, how to abstract phenomena into networks. This includes, for example, what constitutes an individual entity or a relationship, how to conceptualize the strength of a tie, etc. In such applied network science, the corresponding theories are epistemological—network theories bound to specific classes of phenomena. On the other hand, network theories can deal with formalized aspects of network representations such as degree distributions, closure, communities, etc., and how they relate to each other. In such pure network science, the corresponding theories are mathematical—theories of networks.

Claim 2

There are theories about network representations and network theories about phenomena: both constitute network theory.

Establishing network theory can be a challenge in disciplines that have a highly individualized history. But without a theory about how to conceptualize a phenomenon as a network, there is no meaning to a formal theory of network data. We conclude that networks are not just an add-on to existing approaches—e.g., a means to add a little more explained variance in a social science research project—but require new theorizations and di erent thinking.

A network abstraction involves ontological commitment to a few basic features that are seen as scientifically relevant to the representation of a phenomenon. The features of a network abstraction include at least the following: individual elements; pair-wise relationships between those elements; and a global or macro- patterning that can be considered as network structure. This basic description may not be su cient for all circumstances (e.g., think of longitudinal phenomena) and can be extended in many di erent ways, but these are fundamental features if we are to call the abstraction a network.

For example, a friendship network is a way of abstracting a social phenomenon into a comparatively simpler and much more general form of relationships between actors. The actual phenomenon is, of course, much richer: We are abstracting already quite substantially just by conceiving of the individuals as comparable entities of a common kind.

By postulating a friendship network in (say) a school classroom of 25 students, we have taken a theoretical step that is non-trivial. We have supposed that separate individualshapo(e)-315.-387.2(gl—butl2aindividu0actioade)-332

So the claim that "networks are everywhere," if it is meaningful as network science, is not just a statement that we can see many things in the world in relational terms, but an implicit theoretical statement that scientific explanation of many phenomena is aided by abstraction to such a connected, systemic representation. Otherwise, it is no more than a statement that we can see the world in particular ways: after all, colors are everywhere, too, but no one to date has thought it scientifically helpful to understand classroom processes in shades of pink and purple.

The essential point must be that the abstraction into a network is helpful to scientific inference, permitting knowledge to develop. It need not be the case that this will always be so. But we need an empirical base to show that the network representation gives scientific traction.

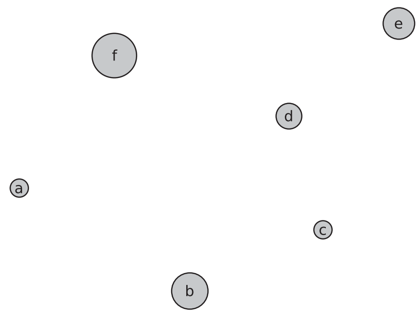
Claim 3
Network science should be empirical—not exclusively so, but consistently—and its value assessed against alternative representations.

3 Network data

We have argued that networks are abstractions represented in data, but we have yet to discriminate them from other conceptualizations. We are now going to do so by first looking at characteristics of standard types of data to be able to then highlight the defining features of network data.

The input to data analysis consists of values of variables. Variables are generic placeholders characterizing the essential features of an abstract concept, thus allowing to formulate analytical steps generically as well. The instantiating values are usually obtained via some form of observation such as measurement. Note, however, that di erent original phenomena may yield the same representation in data.

Our definition of what constitutes network data hinges entirely on how the involved variables are related. It is thus independent of the phenomena being



A	
	1
	4
	1
	2
	3
	6

(a) a a ab : a ab a

In the above example, variables gender, education, and income are of different types: While all are defined on the common domain A, the range of values they may assume is different. Even more importantly, these ranges exhibit a different level of structuring. The range of gender, for instance, is binary on a nominal scale, i.e., the only defined relation is an equality predicate. In other words, the comparison of two values yields either equality or inequality, and this is the only information we can get out of comparison. For instance, we cannot add or rank values of gender variables.

Assume now that education refers to the highest degree obtained by an individual. It may then be valid to compare two values and conclude that one indicates a higher level of education than the other, and this relation could be transitive. In this case, the range is ordered and the variable is on an ordinal scale of education. Finally, we may compare income by amount, but it may also be meaningful to compute differences and ratios. The range of the variable income can therefore be considered to be a continuous ratio scale. If, however, 0 is not meaningful for a continuous range as in, e.g., measuring IQ, it is not appropriate to calculate ratios and the scale is called interval.

The interesting thing to observe is that a range is usually not just a set of possible values but a set with additional relations such as an ordering or operations, i.e., structured. The structure of a range is crucial to know about because it determines the kinds of analyses and interpretations that are justified.

While the range of attributes is structured, in much of science, the domain on which variables are defined is assumed to have no structure, i.e., simply a set. This may be for good reason. If we are interested in associations between, say, education and income controlled for age, we actually do not want there to be relations between individuals that also moderate the association. Much of statistics is in fact concerned with detecting and eliminating such relations.

This is the single most important difference with network science, where the domains of at least some variables are explicitly set up to have structure. The potentially resulting dependencies are not a nuisance but more often than not they constitute the actual research interest.

3.2 Step 2: Dyadic data

Before introducing the structure of network-variable domains, consider as an intermediate step dyadic variables, i.e., variables defined on pairs of items. A classic example of this kind is the study of (populations of) couples. Here variables, such as duration of marriage, number of common children, etc., are associated with the couples, whereas further variables, such as age, occupation, etc., are associated with the individuals that make up the couples.

As illustrated in Figure 2(b), the domain of couple-level variables is therefore composed of pairs of individuals that cannot be treated as a new, atomic unit because it is important to maintain individual identities to be able to find between-variable associations in individual- and couple-level data. Similarly, attributes of couples cannot be represented in individual-level variables because this means eliminating, for instance, possible between-variable associations of individuals and marriages that are moderated by attributes of spouses.

Editorial

because they are good friends who share information relevant for their salary negotiations. In other words, rather than being a dependence between two di erent types of variables, now we have a dependence within the values for one type of variable. This is a complex dependence because it cannot be aggregated or averaged in distributional terms. It corresponds to the kind of dependency analyzed in spatial statistics (with proximity rather than friendship as the underlying mechanism).

Yet, network dependence goes further because dependence does not just stop at actor attribute variables. It may apply within the set of network variables as well. Any network variable is defined on a domain of pairs of individuals (i.e., the dyads), and the incidence structure of the domain captures the potential for within-variable dependencies. A network tie variable takes a value, often binary, sometimes valued, indicating whether there is or not a tie between its two individuals. The crucial point is that the presence of one tie may influence the presence of another. In other words, ties are not necessarily moderating variables, but there may be dependencies within the tie variable themselves. While this will appear an unfamiliar point of view to some, it is merely a statement that networks may be systematically patterned. Without dependence among ties, there is no emergent network structure.

In the explicit form of stochastic models, these ideas entered network analysis from spatial statistics. They are deeply at the heart of network theory, even if seldom overtly addressed. Entire sets of methodological approaches, such as exponential-family random graph models, depend on modeling tie dependence appropriately.

With independence among network tie variables, we would be left only with the simple random networks known as Bernoulli graphs, Erdős-Renyi graphs, or the (n, p) model. It should be noted that this view does not require a statistical perspective; combinatorial invariants of graphs that represent networks are of interest exactly for the same reason as descriptors of structural features.

Because almost all the networks that we observe bear little resemblance to simple random graphs, tie dependence is empirically very common. For instance, a familiar network process is that of preferential attachment, whereby actors “prefer” to be attached to popular actors so that the rich get richer. The presence of many ties centered on one popular individual may attract the presence of additional ties to that same individual.

Dependence among ties is thus the means whereby network structure self-organizes and evolves, or emerges, but it is not simple. This is why network science is often referred to as the study of complex networks. It remains a research question to establish plausible types of tie dependence. Theories or methods that wish away these dependencies are ignorant of the structure of the domain, and thus contradictory to a network model.

While the choice of representation is indeed a matter of convenience and hence

At Network Science, we anticipate to publish work on all kinds of network data, researching in a qualitative or quantitative way, but that the understanding of the phenomenon treats relational connectivity and dependence as central.

4 The emerging science of networks

In light of the above discussion, we hope that this journal will provide a shared intellectual space for network scientists working in many di erent fields to communicate with each other about relational data.

To get there, we must recognize our various shared and disparate histories, recognize that this field is quick-evolving, commit to compatible languages about networks, and be willing to speak outside narrow disciplinary interests to broader communities of scholars. The benefit, we believe, will be well worth the e ort.

As editors of a journal attempting to encompass a broad field with a long and storied history, we have already rejected the idea that network science “began” with some kind of new discovery or even a Kuhnian paradigm shift tipped o by work originating from physics, no matter how interesting or influential. Network science is neither tied to nor “owned” by any other field.

We should not be ignorant of the forebears of our emerging science, and decades of empirical research. The past 15 years have seen a boom of interest in networks that does not overtly trace its roots to, for example, the sociometry of Moreno or the sociology of Simmel. Even this older tradition has long borrowed from other fields such as graph theory, physics, or statistics as it has developed.

Neither are these the sole progenitors of what we now recognize as network science.

This goal transcends disciplinary boundaries but we do have disciplinary goals as well. Our major fields of editorial coverage (with area editors in parentheses) include information science (Adamic), computer science and mathematics (Brandes), communication, engineering and management (Contractor), economics (Goyal), political science and psychology (Robins), public health and medicine (Valente), physics (Vespignani), and statistics and sociology (Wasserman). Each editor has identified key topics and debates within their area that they would like to see addressed in the coming issues of Network Science and that list follows this editorial as an attachment. Consider these an open call for work, but also consider Network Science as welcoming of work that pushes this new science forward.

We are excited by the prospects of this new journal, Network Science. We believe there is a distinctive science of networks that crosses traditional disciplinary boundaries. It is ready to be brought together in a coherent form that transcends disciplinary silos. We encourage all our readers to contribute to the journal to help achieve these goals.

Acknowledgments

Part of the research that led to this editorial was funded by Deutsche Forschungsgemeinschaft under grant Br 2158/6-1 and the Social and Cognitive Networks Academic Research Center of the US Army Research Laboratory.

Notes from the area editors

Lada Adamic, editor for information science

Information is an interdisciplinary field, just as network science. Therefore, a broad range of topics can fall under this heading, including networked information (e.g., the web, Wikipedia, citation networks), information dynamics in online, organizational, and other social networks, and networks that can be constructed by representing relationships between data (e.g., health, scientific, or historical data). We invite contributions that include novel theoretical models, empirical studies, and methods and applications pertaining to information networks.

Ulrik Brandes, editor for computer science and mathematics

We invite articles presenting original research in structural and computational network science. This includes the study of network representations, algorithms, data management, and visualization. A typical theory paper uses graph theory, combinatorics, algorithmics, machine learning, information retrieval, or computer graphics methods, whereas a systems paper concentrates on design aspects, implementation, and performance assessment. Novel uses of network approaches in application areas, and in particular those relating to social media, may also be suitable for the information science area.

and interaction of individual attitudes, traits and behaviors, and social network ties, including network-based social influence. Finally, we are also interested in the