December 2014 | Volume 102 | Number 12

Proceedings of the IEEE

SPECIAL ISSUE

Impact of Changing Technology on Social Networks

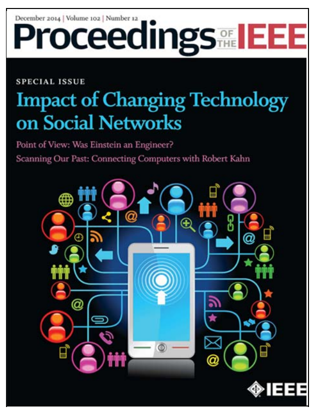
Point of View: Was Einstein an Engineer?

Scanning Our Past: Connecting Computers with Robert Kahn

IEEE

Proceedings of the IEEE
DECEMBER 2014 / VOL. 102 / NO. 12

CONTENTS



On the Cover: Our cover image this month captures the essence of social networks and changing technology in a visually appealing graphical form.

The Digital Age and the Future of Social Network Science and Engineering

By JESSICA C. FLACK
Guest Editor

RAISSA M. D'SOUZA
Guest Editor

This special issue of the PROCEEDINGS OF THE IEEE focuses on how digital technology is changing the structure and dynamics of social networks and the tools we have for studying and designing them. As a testament to the importance of this topic consider that as of 2014, 72% of Internet users in the United States and 64% of users worldwide use social media [1], [2]. Facebook alone has 1.15 billion users, up from 1 million in 2004, just ten years ago [2]. In the United States, the average user spends more than one fourth of every online hour on social media, and almost 50% of Americans say that Facebook is their #1 influencer of purchases [2]. Google+ has been around for only three years and already there are 1 billion Google+ enabled accounts [2].

This special issue focuses on how digital technology is changing the structure and dynamics of social networks and the tools we have for studying and designing them.

This rapid growth in connectivity stands in stark contrast to the vast majority of human history in which social networks were small and geographically localized, institutions changed slowly, and power and influence were concentrated in a small subset of the population [3]. The rise of the world wide web—and particularly the invention of powerful search technologies, social media platforms, and novel file sharing technologies—has led to an explosive growth in social network size and connectivity as well as the development of new kinds of reputation-based barter systems (e.g., reviewed in [4] and [5]) and underground economies (e.g., [6]). Standard geographic definitions of population compete for causal relevance with definitions that group people together based on behavioral criteria, including hyperlinking on the web [7]. It is now possible to track events as they unfold in real time, to coordinate relatively rapidly over long space scales (for a potential example, see [8]), and to access semiglobal and global information to make decisions. This means that individuals and subgroups without much power in the traditional sense potentially can serve as instruments of large-scale social change. The growth of social networks on digital media also means it is possible to collect big, reliable data sets on human behavior and associated events such as earthquakes [9], as human behavior on digital media leaves a relatively easily harvestable data trail [10].

These technological changes, the new social structures they seem to be producing, and the data being generated permit reconstructing the behavior of individuals from their digital footprints (e.g., [11]; see also the paper by Coviello *et al.* of this special issue [12]). This will enable empirical study of *how* different kinds of social structures arise from interactions among individuals at the microscopic level and *what* the implications are of alternative social structures for social stability and contagion processes. As progress is made on this front, large-scale intervention into human

Digital Object Identifier: 10.1109/JPROC.2014.2368790

behavior and demographics will no longer be restricted to coarse manipulation of environmental variables or interventions based on qualitative understanding of how social systems work (as an example, consider the so-called causal loop diagram produced to describe the counter insurgency dynamics in Afghanistan [13]).

Precision, quantitatively justified interventions into behavioral dynamics will become increasingly feasible. Already such interventions are being employed; the Facebook "61-million-person experiment" in which users' affective states were manipulated [14] serves as one example. These interventions will allow the interveners to influence individual decision-making and permit modulation and possibly control of macroscopic properties of social systems; currently, for example, large, digital data sets on online social behavior are being harvested to inform the design of social media platforms (e.g., [15]). Although these large-scale experiments on human behavior will initially occur in the digital world, they will eventually inform the mechanism design of real-world infrastructure to optimize communication and better meet basic human needs such as access to food, sanitation, healthcare, and schooling (e.g., [16]).

Hence social media, search, and data extraction technologies are not only changing the structure and dynamics of social networks but are also potentially changing how *controllable* these systems are. The purpose of this special issue is to review what is currently known about how these technologies are changing social networks and what the consequences will be for human social dynamics. Understanding the impact of changing technology on social networks is relevant to the readers of the Proceedings of the IEEE for several reasons. One reason is that it is engineers, mathematicians, and computer scientists who are largely responsible for developing the social media technology, file sharing protocols, and pattern extraction algorithms that are producing changes to social networks. An understanding of cognitive principles and behavioral interaction rules underlying social network formation can facilitate the development of social media applications that individuals are more likely to use.

A second reason is that an understanding of cognitive and behavioral constraints could improve the performance of machine learning and other pattern detection and data extraction algorithms by providing principled means to restrict the search space.

Finally, with the rise of online social networks information becomes a concept as central as energy to informing design of infrastructure. Societies are composed of multiple, overlapping social networks. This is true for networks developing in "real" space as well as on digital media. The local connections of a node in these multiple, overlapping social networks can be thought of as that node's social niche [17], with the edges in each of the networks representing different kinds of relationships, for example, friends, co-workers, mentors, and students. Whereas the ecological niche [18] (the term "niche construction" was invented in ecology; see [19]) is composed of resource vectors (availability of wood for building dams, prey items, and so on), an individual's social niche is composed of its vector of behavioral relationships in the set of overlapping social networks in which it participates [17]. These behavioral relationships provide critical information that facilitates resource extraction, and the construction of these networks by users can be thought of as a form of social niche construction [17], [20].

To the extent this conceptualization is correct, social networks—perhaps more than other types of networks—are about the flow of information. This means that higher order features of network structure (that is, minimally, connections beyond a node's direct connections) may be of critical impotence to understanding the functional consequences of a given network for its nodes, both individually and collectively. This observation is increasingly valid as social networks move online and become further divorced from direct resource extraction. In addition, issues of sociotechnical congruence—the match or mismatch between social networks and the technological networks they build and rely upon—have implications for information flow that need to be considered when assessing how information impacts function and energy extraction (e.g., [21]). As Weaver and Shannon [22] noted more than 60 years ago, although we have a good theory of information defined as uncertainty, we have no theory relating information to function. The migration of social interactions into the digital domain means that the absence of semantic and functional theories of information is no longer a curious omission but a central challenge for theory and analysis.

In the first paper, Bettencourt suggests that a property of social, informational networks is superlinear scaling, which translates into increasing returns to scale. Bettencourt proposes that network dynamics in these kinds of complex social systems need to be understood in relation to evolutionary learning and inference principles. He develops a conceptual framework to explain the transition from relatively static, homogenous, information-poor networks to information-rich, diverse, and highly interconnected ones using current understanding of social scaling in urban environments to develop his argument, and then extending it to the digital social domain using the Internet and Wikipedia as case studies.

Networks typically have organization at multiple levels of scale, from correlations between individual nodes to larger scale modular structures and groupings into communities of nodes with similar characteristics. Understanding how to reliably detect these structures remains a central activity in the study of networks, especially as the goal is to connect structural features with the function of a network. This task is increasingly challenging as we comprehend the extent to which multiple communities can overlap in nontrivial ways. In the second paper, Yang and Leskovec tackle this issue by considering each community as a "tile" which can overlap extensively with other tiles. Their methodology decomposes the network into a combination of overlapping, nonoverlapping, and hierarchically organized communities. In contrast to previous work, they show that nodes residing in the overlap between communities are more densely connected than those in nonoverlapping regions. Moreover, they show that overlap between multiple communities identifies dense network cores, revealing the core–periphery structure in networks and thus offering a method to unify the study of network modularity and core–periphery structure.

Singh *et al.* point out, in the third paper, that the line between real or physical social networks and those on digital media is not as strict as many assume. The authors observe that fine-grained behavioral data on individuals collected from online sources and smartphones are now being merged to give an increasingly comprehensive picture of an individual's location, social ties, actions, and context in both settings. These merged data can be used to influence the actions and beliefs of individuals in both settings, largely because smartphones function as a bridge between worlds. This bridge allows individuals to behave simultaneously in both the digital and physical domains, createing real-time couplings between their social behavior in both worlds.

The question of how coupling occurs between social behavior in the online and physical worlds can be reframed as a question of synchrony and contagion: How does behavior spread over social networks through contagion and other mechanisms and when do contagion dynamics lead to correlation in social preferences and decision making and even behavioral synchrony? The large quantity of quantitative data on behavior and contagion available from social media sites means that this question can be addressed at a variety of scales, and it means that, in principle, the causal mechanisms and factors that influence spread and synchrony can be identified.

There are two broad classes of approaches that can be used to quantify causal relationships. One is experimental intervention. This can include perturbing the social media platform itself (e.g., how individuals are allowed to make posts, how long those posts can be, etc.), perturbing the social media environment (e.g., changing the aesthetics of the site, site branding, types of advertising to which users are exposed, etc.), and injecting information or behavior into the system (e.g., creating users and posts), and tracking its effects using time-series analysis.

The second approach for quantifying casual relationships is observational with statistical "interventions" (e.g., Pearl's "do" operator [23]). Although it is relatively easy to perform large-scale experiments using social media data, there are, as the Facebook controversy discussed above illustrates, many ethical issues that need to be carefully considered before proceeding. Developing methods for performing "statistical" interventions on observational data is consequently of utmost importance. The fourth paper by Coviello *et al.* develops a method for quantifying influence of users on one another. The method combines geographic aggregation and instrumental variables regression to measure the effect of an exogenous variable on an individual's expression and the influence of this change on the expression of others to whom that individual is socially connected. The authors demonstrate the power of the approach in the context of emotional contagion of semantic expression but also show that the approach is quite general and can be applied to many kinds of data collected from a variety of social networking platforms.

These ideas are reviewed in a broader, technical context by Holme, in the fifth paper. Holme highlights that the data available from social media sites and online social networks typically have time stamps. These time stamps allow analysis of *how* the timescales on which contacts and interactions occur can influence contagion dynamics and network construction. Holme reviews methods for analyzing temporal networks. He addresses challenges in representing temporal interactions in networks and in the construction of appropriate null models of temporally evolving network structure. He also discusses principled means for simplifying temporal networks to make them amenable to rigorous analysis.

As discussed above, a growing proportion of human activity is leaving behind digital footprints. In the sixth paper, Lamboitte and Kosinski show how pervasive details left in these footprints can be used to infer an individual's personality, where personality refers to the major psychological framework identifying individual differences among people in behavior patterns, cognition, and emotion. In addition to reviewing the factors of personality, they review a range of recent works focused on predicting personality from digital traces. They conclude with discussion of the increasingly important implications for privacy, security, and data ownership as well as future directions for research.

In the final paper, Walker and Muchnik delve into how our traditional methods of experimental design, which divide subjects into control and treatment groups, need to be radically redesigned for our digital world. Traditional methods neglect the natural connections that exist between individuals, and, more so, that the impact of treatment can propagate through such connections. In analogy to big data, they survey the burgeoning movement of big experiments (such as the 61-million-person experiment on Facebook [14]). They discuss the broad range of aspects that need to be considered for appropriate design of network-randomized trials, including the experimental setting, the process under study, and the impact of connectivity. Equally important is their discussion on emerging methods to draw statistically meaningful conclusions from such experiments as well as developing novel treatment schemes that leverage connections in social networks. They repeatedly highlight the important and subtle distinctions between offline and online experimental settings.

In conclusion, this is an exciting moment in human history as the digital world and the physical world become increasingly intertwined in a seamless manner. We offer three take-home messages.

- Social media, search, and data extraction technologies are not only changing the structure and dynamics of social networks, but are also changing how controllable these systems are.
- Precision, quantitatively justified interventions into behavioral dynamics are increasingly feasible within the digital domain, permitting large-scale experiments on human behavior and social systems. This is useful and presents challenges.
- We understand the relationship between energy and information—how bits get converted to watts—for electrical circuits, but not for social networks. In biology, computational social science, and the science of social engineering, the development of a functional theory of information is a central theoretical challenge that needs to be addressed if these disciplines are to have strong foundations. ■

Acknowledgment

The guest editors would like to thank D. Krakauer and G. Spencer for comments and ideas for the content and structure of this special issue. The work of J. C. Flack was supported by a grant from the John Templeton Foundation for the study of complexity. In addition, the work of J. C. Flack and R. M. D'Souza was supported by the U.S. Army Research Office Multidisciplinary University Research Initiative (MURI) Award under Contract W911NF-13-1-0340.

REFERENCES

[1] Statista, "Number of social network users worldwide from 2010 to 2017 (in billions)." [Online]. Available: http://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/

[2] J. Bullas, "22 social media facts and statistics you should know in 2014." [Online]. Available: http://www.jeffbullas.com/2014/01/17/20-social-media-facts-and-statistics-you-should-know-in-2014/

[3] J. M. Diamond, *Guns, Germs, Steel: The Fates of Human Society*. London, U.K.: Norton, 1999.

[4] A. Josang, R. Ismail, and C. Boyd, "A survey of trust and reputation systems for online provision," *Decision Support Syst.*, vol. 43, pp. 618–644, 2007.

[5] P. Resnick, K. Kuwabara, R. Zeckhauer, and E. Friedman, "Reputation systems," *Commun. ACM,* vol. 43, pp. 45–48, 2000.

[6] A. Greenberg, "Drug market 'agora' replaces the silk road as king of the dark net," *Wired*, Sep. 2, 2014. [Online]. Available: http://www.wired.com/2014/09/agora-bigger-than-silk-road/

[7] D. Gibson, J. Kleinberg, and P. Raghavan, "Inferring web communities from link topology," in *Proc. 9th ACM Conf. Hypertext Hypermedia: Links, Objects, Time and Space—Structure in Hypermedia Systems*, 1998, pp. 225–234.

[8] W. Ghonim, *Revolution 2.0: The Power of the People Is Greater Than the People in Power*. New York, NY, USA: Houghton Mifflin Harcourt, 2012.

[9] L. Burks, M. Miller, and R. Zadeh, "Rapid estimate of ground shaking intensity by combining simple earthquake characteristics with tweets," presented at the 10th U.S. Nat. Conf. Earthquake Eng., Front. Earthquake Eng., Anchorage, AK, USA, Jul. 21–25, 2014, 10NCEE.

[10] Edge, "Reinventing society in the wake of big data: A conversation with Alex (Sandy) Pentland," Aug. 30, 2012. [Online]. Available: http://edge.org/conversation/reinventing-society-in-the-wake-of-big-data

[11] N. Eagle, A. Pentland, and D. Lazer, "Inferring social network structure using mobile phone data," *Proc. Nat. Acad. Sci.*, vol. 106, no. 36, pp. 15 274–15 278, 2009.

[12] L. Coviello, J. H. Fowler, and M. Franceschetti, "Words on the web: Noninvasive detection of emotional contagion in online social networks," *Proc. IEEE,* vol. 102, no. 12, Dec. 2014, DOI: 10.1109/JPROC.2014.2366052.

[13] PA Consulting Group, "Dynamic planning for counterinsurgency in Afghanistan, PowerPoint presentation," 2009. [Online]. Available: http://msnbcmedia.msn.com/i/MSNBC/Components/Photo/_new/Afghani-stan_Dynamic_Planning.pdf

[14] R. M. Bond*et al.* "A 61-million-person experiment in social influence and political mobilization," *Nature,* vol. 489, pp. 295–298, 2014.

[15] V. K. Singh, R. Jain, and M. Kankanhalli, Mechanism design for incentivizing social media contributions," in *Social Media and Modeling*, Hoi*et al.*, Ed. London, U.K.: Springer-Verlag, 2011.

[16] A. Wesolowski*et al.* "Quantifying the impact of human mobility on malaria," *Science,* vol. 338, pp. 267–2702014.

[17] J. C. Flack, M. Girvan, F. B. M. de Waal, and D. C. Krakauer, "Policing stabilizes construction of social niches in primates," *Nature*, vol. 439, pp. 426–429, 2006, see supplementary information for social niche definition.

[18] G. E. Hutchinson, "Concluding remarks," *Cold Spring Harbor Symp. Quantitative Biol.*, vol. 22, pp. 415–427, 1957.

[19] J. F. Odling-Smee, K. N. Laland, and M. W. Feldman, *Niche Construction: The Neglected Process in Evolution. Monographs in Population Biology 37*. Princeton, NJ, USA: Princeton Univ. Press, 2003.

[20] J. C. Flack, "Multiple timescales and the developmental dynamics of social systems," *Phil. Trans. Roy. Soc. B 5,* vol. 367, pp. 1802–1810, 2012.

[21] D. Posnett, R. M. D'Souza, P. Devanbu, and V. Filkov, "Dual ecological measures of focus in software development," in *Proc. 35th Int. Conf. Softw. Eng.*, 2013, pp. 452–461.

[22] W. Weaver and C. E. Shannon, *The Mathematical Theory of Communication*. Urbana, IL, USA: Univ. Illinois Press, 1949.

[23] J. Pearl, *Causality: Models, Reasoning, Inference*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

ABOUT THE GUEST EDITORS

Jessica C. Flack received the B.A. degree with honors from Cornell University, Ithaca, NY, USA, studying anthropology, evolutionary theory, and biology, and the Ph.D. degree from Emory University, Atlanta, GA, USA, studying animal behavior, cognitive science, and evolutionary theory.

She is a Codirector of the Center for Complexity and Collective Computation (C4), Wisconsin Institute for Discovery, University of Wisconsin at Madison, Madison, WI, USA and External Professor at the Santa Fe Institute, Santa Fe, NM, USA. After her Ph.D., for the next eight years, she was in residence at the Santa Fe Institute, first as a Postdoctoral Fellow and then as Professor. She moved to the University of Wisconsin at Madison in 2011 to build C4. Her research focuses on coarse graining and collective computation in nature and their role in the evolution and development of new levels of biological and social organization. The approach is highly empirical, beginning with rich microscopic data, including network time series, gained from model systems in neuroscience, animal behavior, and human social systems. The goal is to find common algorithmic principles underlying the emergence of novel, functionally significant spatial and temporal scales, and ultimately new kinds of collectives and individuality. This research has involved development of computational techniques (*Inductive Game Theory*) for extracting strategies (and potentially game structure and payoffs) from time-series data and constructing stochastic strategic circuits that map input to the target macroscopic output. Dimension reduction techniques are used to identify the most appropriate minimal model for representing the compressed circuits. One can think of this approach as a *strategic statistical mechanics*, embedding complex decision rules in formalisms for connecting microscopic configurations to macroscopic averages. An ultimate goal is to determine whether living systems—composed of noisy, adaptive, heterogeneous components often with multiple goal states—are governed by laws derived from simple microscopic processes or reflect contingent events leading to irreducible complexity. Her work has appeared in journals such as *Nature*, *Proceedings of the National Academy of Sciences of the United States of America*, *Royal Society Interface*, and *Chaos*, and has been covered by scientists and journalists in popular press outlets, including *Science*, *Nature*, *The Economist*, *Current Biology*, and the *BBC*.

Raissa M. D'Souza received the Ph.D. degree in statistical physics from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 1999.

She is a Professor of Computer Science and of Mechanical Engineering at the University of California Davis, Davis, CA, USA, as well as an External Professor at the Santa Fe Institute, Santa Fe, NM, USA. She was a Postdoctoral Fellow, first in Fundamental Mathematics and Theoretical Physics at Bell Laboratories, and then in the Theory Group at Microsoft Research. Her interdisciplinary work on network theory spans the fields of statistical physics, theoretical computer science, and applied math, and has appeared in journals such as *Science*, *Proceedings of the National Academy of Sciences of the United States of America*, and *Physical Review Letters*. Her research has been funded by the National Science Foundation (NSF), the National Academy of Sciences of the USA, the U.S. Department of Defense (DOD), the U.S. Army Research Laboratory (ARL), and the U.S. Air Force Office of Scientific Research (AFOSR). Her research focuses on mathematical models of complex systems and networks, in particular on self-organization, phase transitions, and the structure and function of networked systems. Interdependence between collections of networks and sociotechnical congruence are additional themes underlying her research.

Prof. D'Souza has been selected a Kavli Fellow of the National Academy of Sciences four times and is a member of the Science Steering Committee of the Santa Fe Institute. She has served on the program committees for numerous conferences, including serving as the General Chair of NetSci 2014, and currently serves on the editorial board of numerous international mathematics, physics, and engineering journals. She is also a member of the World Economic Forum's Global Agenda Council on Complex Systems.

Impact of Changing Technology on the Evolution of Complex Informational Networks

This paper proposes a general framework to understand the fundamental advantages of connectivity in complex informational networks.

By Luís M. A. Bettencourt

ABSTRACT | We live in an era of increasing connectivity in human societies and in technology. These structural changes in the ways we interact with each other and with increasingly ubiquitous computational and communication devices have been formalized in research across several disciplines through the dynamics of complex informational networks. Complex networks are (mathematical) graphs, connecting nodes (people, computers) via edges (relationships, wires). While much progress in methods for network analysis has been achieved, the fundamental principles that drive network growth in human societies and in worldwide computer networks remain rather obscure. Mechanistic models for the origin of certain structural graph elements have now become common, but the formal connection between large empirical studies of network evolution and fundamental concepts of information, learning, and social theory remains only latent. To address these issues, I argue here that the most interesting aspect of the dynamics of informational networks in complex systems is that they are the physical manifestations of processes of evolution, inference, and learning, from natural ecosystems, to cities and to online environments. I formalize the general problem of learning and computation in network environments in terms of average structural network changes and propose a conceptual framework to explain the transition from initially static, undifferentiated, and information-poor environments to dynamical, richly diverse, and interconnected systems. I illustrate these ideas empirically by providing examples from cities, and from global computer networks and webs of documents. I finish with an overview of expected changes to urban form and function and to computational hardware under likely technological scenarios.

KEYWORDS | Collaborative work; communication networks; complex networks; distributed computing; intelligent systems; learning systems; urban areas; Wikipedia; World Wide Web (WWW)

Manuscript received May 10, 2014; revised October 15, 2014; accepted October 27, 2014. Date of current version November 18, 2014. This work was supported in part by the Bill and Melinda Gates Foundation under Grant OPP1076282; the MacArthur Foundation under Grant 13-105749-000-USP; the James S. McDonnell Foundation under Grant 220020195; the John Templeton Foundation under Grant 15705; the Army Research Office Minerva program under Grant W911NF1210097; and by a gift from Jeannie and Michael L. Klein.

The author is with Santa Fe Institute, Santa Fe, NM 87505 USA (e-mail: bettencourt@santafe.edu).

Digital Object Identifier: 10.1109/JPROC.2014.2367132

I. INTRODUCTION

We live at a time of increasing connectivity. This is true of many of our most important technologies as well as of human societies themselves [1]. The rise of telecommunications and of globally networked information technologies, such as the World Wide Web (WWW), is clearly changing human societies in many ways, from our ability to create and store information and run civic institutions [2], [3] to scientific research [4], and from technological innovation [5] to human development [6], [7]. Not only are people and human organizations increasingly connected worldwide but also so are devices and engineering systems, through developments in information and communication technologies (ICTs), such as the Internet of Things [8].

Why do these trends toward greater global connectivity seem so irresistible? Why now? To attempt to answer these questions, I propose that we must build a general framework to understand the fundamental advantages of connectivity in complex informational networks, as well as their associated costs. Through the analysis of these tradeoffs and their change over time we will gain new perspectives from which we can take a new unified look at the history of many technologies—from cities, to transportation and telecommunications—and assess the future of a more connected global human society and its relationships to Earth's other complex adaptive systems [9].

The ideas developed below are about informational networks, which are a subset of what we now refer to as complex networks [10]. Informational networks are made of nodes that are capable of changing their internal states in ways that can expand their information content. This happens in tandem with changes of connectivity as determined by their relative benefits and costs. For example, people can acquire new knowledge and expertise and exchange it through an expanded set of socioeconomic relations with others. In this sense, I will not have much to say about complex networks such as power grids, metabolic graphs, or transportation networks, which do not share these properties. Examples of informational networks may include human socioeconomic systems at different scales, networks of evolving documents, as well as, possibly, ecological networks and neuronal networks capable of learning (natural and artificial), though the latter will not be discussed here.

Because they deal with learning, informational networks pose a set of difficult problems, tied to the dynamics of innovation and productivity in distributed systems. How then may we start to understand these systems in general but simple ways? I propose here that the place to start is the realization that general computation and network connectivity are useful, because information, its discovery, and aggregation are necessary conditions for development in human societies [6], [7], [11] and for learning processes more generally. This point is not new: It is the basis for our best current ideas about economic growth [12] and human development, as well as, in different forms, about natural and technological "evolution" [5], [13], [14].

The interplay between the structure of various complex networks and their embedded information runs in both directions. While technological change that enables larger and more connected individuals provides the conditions for differentiation and learning [15]–[17], it is also true that the acquisition and management of information will remain limited when not embedded in dynamical network structures [18].

Specifically, the creation and management of information typically requires greater human social connectivity through the specialization and interdependence of knowledge in individuals and organizations [19]. These processes of individual and social change are much older than technological progress in modern computing. Although some precedents existed in simple human societies [20], they became manifest in earnest with the advent of the first urban civilizations [21]. With the industrial revolution and the development of new transportation and communication technologies, these processes gained new speed and scope, culminating with the worldwide information revolution currently under way [1], [22]. Thus, it is important to understand the basic conditions necessary for networks of intelligent nodes (such as people) to acquire information in open-ended ways, and the nature of situations when such dynamics may stall. Providing a simple (mean-field) framework to understand these dynamical changes is the main objective of this paper.

II. INFORMATIONAL BASIS FOR INCREASING CONNECTIVITY

Unlike what happens in simple physical systems where interactions between a system's elements constrain its overall structure [23], increases in connectivity in informational networks can, in some specific circumstances, lead to greater individual freedom. This statement may appear paradoxical and requires further explanation. In this section, I formalize these ideas and show when a dynamics of diversification and learning in networks can develop and how it benefits from technological change.

Before I introduce a more formal description, I would like to tap on some of our common intuition for this familiar but perhaps unexpected phenomenon. Many human social environments encourage, and indeed require, that individuals pursue different interests and vocations [24]–[26]. In modern human societies, these processes are partly formalized in educational and professional organizations. In this sense, individuals are encouraged to learn and create new information and new expertise and should expect to be rewarded for such efforts. Much of the modern explosion in technology, science, and the arts depends on the properties of these environments [24]–[26] as does entrepreneurship and human development [6], [7]. Urban [21], [24]–[26] and online environments [27] are general examples of networks where such dynamics of personal expression, learning, and sharing are not only possible on vast scales, but are, in fact, in some sense necessary.

Thus, the type of freedom that networked systems open up is intimately tied to their ability to allow individuals and groups to acquire new information, new roles, and new relationships in fluid ways. In other words, this freedom resides in the various forms of knowledge and relationships that nodes can create, and not so much on the reduction of the number of such interactions [26], [28].

Crucially, this increase in diversity and (implicitly) in individual learning and social expression in networked systems has important material consequences to human societies as it typically leads to greater (economic) productivity as a function of the size of the system, a concept know in economics as increasing returns to scale [29]. Below I discuss in more detail how this can happen as the result of network effects and, in fact, how these effects not only follow, but are necessary, to support a process of increasing connectivity and learning as part of a dynamical virtuous cycle of development. Important examples of these effects are that the economies of bigger cities tend to be larger on a per capita basis [30]–[32] than those of smaller places. Similarly, online systems such as the WWW or Wikipedia become more productive per capita (in ways that I will specify in Sections III-B and III-C) as a growing function of their size.

The essential property for this type of individual differentiation to occur is that nodes can increase their information through learning. Note that this is a necessary but not sufficient condition, as I demonstrate below. This is a natural property of human social networks, but it is currently only incipient in many technological networks, where learning at the node level (webpage, article, computer program) still requires human intervention [33].

The rest of this section is dedicated to formalizing these ideas in general terms, first by grounding the benefits of connectivity in foundational concepts from economics and social sciences, then by discussing the origin and costs of connectivity and finally by bringing it all together in a general quantitative framework for open-ended network dynamics.

A. General Advantages of Connectivity in Informational Networks

Evidence for the advantages of connectivity in informational networks is everywhere, from global trade to living in cities and from the achievements of science to the many uses of the WWW. But it has not always been like this and for a good reason: as I show below, connectivity is very costly and requires system scale and density to pay off. Here, I start from the very beginning by revisiting the basic concepts by which large-scale social connectivity in human societies has been justified and understood. In this way, I attempt to create a picture of how connectivity gradually develops in networks of certain types and how it can become more and more pervasive, under certain conditions that rely on technological change.

The general advantages of connectivity in networked informational systems are perhaps best introduced through the foundational concept of economics [34] and sociology [35], [36]: The division and coordination of labor. In the book that created modern economics, Adam Smith dedicated the three opening chapters of the *Wealth of Nations* [34] to the spectacular increases in productivity achieved by the division of labor. Smith illustrated this point through the increases in productivity of workers at a pin factory: By specializing in different small tasks and coordinating their labor as a whole to produce the final product, Smith estimated that each worker was able to produce about 480 pins/day, while a person working alone may master just a few. Thus, we obtain an increase in average labor productivity of about 100.

What are the basic ingredients of this spectacular gain in labor productivity? Adam Smith originally identified three types of effect, each of which remains important in modern complex networks though they may not be sufficient.

First, he considered the effect of learning to perform a task better, that is, the process of acquiring knowledge and expertise through accumulated experience [37]. This sort of effect has since been extensively studied in manufacturing at the organizational level and in cognitive science, at the individual level [38], [39]. A second source of productivity gains arises from the time savings resulting from avoiding switching between tasks. Finally, a third source of gains relies on the possibility that a task that has been rendered sufficiently simple through the successive division of labor can be made automatic, and, in that sense, be performed by a machine, thus saving human labor. Many technologies started this way, by the observation from a specialized worker of a solution that can save his/her labor and time.

These different sources of productivity gains are very general and clearly transcend the context of economic production in manufacturing. Thus, we should not think of the process of the division of labor, in terms of vertical integration of minutely specialized jobs in manufacturing firms (though it is that as well) but rather of the distribution of tasks in networks that are generally not hierarchies and the necessary creation of knowledge entailed by the specialized task and its integration (recombination) in many products and services [5], [17]. In this form, the ancient concept of the division and coordination of labor gains new life as a modern process, at play around us everywhere. In this modern form, it emphasizes information and communication in evolving complex networks. Many of the most modern socioeconomic phenomena— from online collective intelligence to the share economy—depend in fundamental ways on these processes.

In this light, the creation and interdependence of knowledge requires the development of complex and dynamical network structures as an evolving process that is sketched in Fig. 1. I start, for simplicity, by imagining a situation where a set of nodes (people, or other informational objects, books, computers) each approximately replicates the same functions [denoted by different colors in Fig. 1(a)]. An example is that of a subsistence human society [20], where despite some specialization of labor by sex and age, all households perform essentially the same tasks of hunting, gathering, or small-scale farming. The information content of such societies is replicated in each nuclear family [Fig. 1(a)] as they survive as individual units in interaction with their natural environment. Thus, the total information content in this situation is that of the typical unit, because each node is redundant (nondifferentiated) with all the others. This is why, in this type of disconnected network phase, information does not accumulate with increases in system size (nodes).

The situation changes radically as large-scale connectivity becomes common [Fig. 1(b)] [17]. It is then possible for nodes to differentiate and specialize on different tasks, relying on their functional complementarities to preserve overall function at the network level. So, for example, in modern urban societies most of us do not grow our own food or harvest energy and instead devote our time to extremely specialized tasks, often in services and in learning and organization. We rely on a vast number of different people (most of them strangers) for our survival in terms of

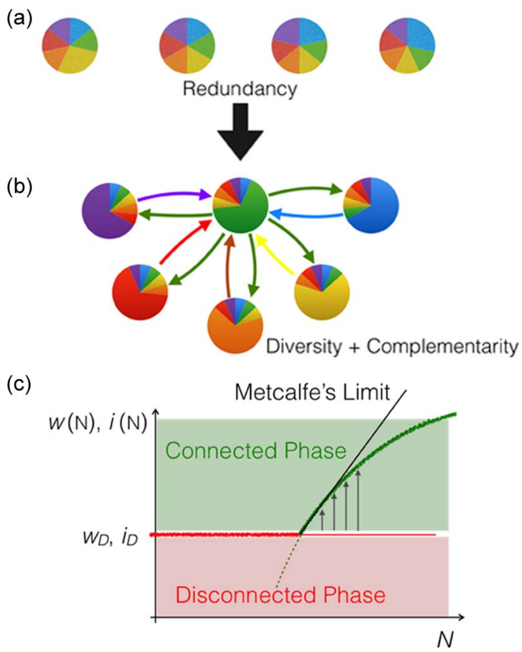


Fig. 1. ***Structural transformation in informational networks resulting from interconnection and knowledge specialization. The disconnected phase (a) is characterized by low levels of connectivity, functional redundancy (duller colors), low productivity, and slow learning; see Section II-B. As connectivity becomes less costly, a transition to a new phase (b) is possible, characterized by increasing connectivity with scale, accompanied by greater complementarity of functions (brighter colors), growing productivity, and fast overall learning at the individual and network levels. In this phase, nodes become functionally interdependent and exchange information, goods, and services (curved arrows). (Only the arrows in which the green node participates are shown, for simplicity.) (c) Nature of the transition (see Section II-D): Small systems where communication and exchange are costly will tend to be in the disconnected phase, while larger systems with inexpensive connectivity will tend to be in the connected phase and under continuous growth of information and productivity with network size N. Thus, as size or technological circumstances change, disconnected systems may become susceptible to entering the connected phase and vice versa. As this happens, the transition may be gradual, along the thick line, or sudden (arrows). Metcalfe's limit refers to the situation when productivity and information increase linearly with system size; see Table 1.***

the most critical products and services we need daily (food, water, etc.). In this situation, the information content of a network can scale up with the size of the society N as individual differentiation becomes the norm. Thus, even if the size of a network N were to remain constant, its information content will now be much larger, roughly proportional to the number of its nodes, naturally conferring economic and technological advantages to large connected systems. This effect is observable in some specific networks, such as associate memory models like Hopfield networks where it can be derived formally [40]. In real human societies, differentiation is typically not fully extensive (many people perform the same professions, for example). To deal with this issue, Adam Smith originally posited that (economic) specialization is in fact proportional to the "extent of the market" [16], [17], [34]. The factors that may limit functional differentiation of nodes in a network and consequently its potential information capacity, productivity, and diversity will be discussed further below together with a simple quantitative framework that makes more precise the qualitative picture introduced in this section.

B. Quantifying the Benefits of Connectivity

Having stated the general benefits of increasing network connectivity in terms of gains in information and productivity, I now quantify these effects in general terms, as functions of network size. I develop a simple "mean-field" model of these processes, where only average properties are taken into consideration. The development of a full statistical model, capable of accounting for the effect of fluctuations on the transition between network phases, requires additional technical development and will be pursued elsewhere.

Let us begin with the simpler disconnected phase; see Fig. 1(a). In this regime, the system is very simple and can be characterized by quantities that are independent of network size, because the nodes lack large-scale connectivity. As such we denote the constant connectivity per node (degree) $k = k_D$, the constant information content $i = i_D$, and the constant productivity per person $w = w_D$; see Table 1. Two related quantities are also worth specifying: the number of functions per person (a measure of individual specialization, e.g., the number of professions per person) $d = d_D$ and the average time spent on each task $t = t_D = T/d$, where T is the total activity time for an individual. As discussed above, I assume that in this phase, each node is a "subsistence generalist," defined by low connectivity and a large number of tasks it needs to perform to survive. As a consequence, the time per task is small, leading to low productivity as each task is associated with relatively small amounts of learning. Though the total knowledge (information) of each node may be large, it is redundant with other nodes performing similar tasks, and the information content of the system is low and of order i_D. In other words, because of low connectivity, individual functional differentiation is minimal and learning (information acquisition) is very slow, as a result of a very small amount of time spent on each task, even if network size N is large. Fig. 1(c) summarizes this situation (constant i_D, w_D on N) by the horizontal red line.

When connectivity and interdependence become possible, a network can express very different properties for the same size N. Let us for a moment ignore the costs of creating and maintaining connectivity, which are discussed in Section II-C. Then, let us suppose further that connectivity per capita increases with N. For illustration, I assume that it varies according to a scale invariant function (a power law) of the form $k(N) = k_C N^{\delta}$, where the amplitude k_C depends on technology and time (and cost)

Table 1 Characteristics of Complex Informational Network Phases

Symbol	Node Property	Disconnected Phase		Connected Phase (general)		Connected Phase (example)	
k	connectivity (degree)	k_D	(small)	$k(N)$	(increasing)	$k(N) = k_C(t)N^d$	(increasing)
d	number of functions	d_D	(large)	$d(N) = Ak^{-1}(N)$	(decreasing)	$d(N) = d_C(t)N^{-d}$	(decreasing)
i	information	i_D	(small)	$i(N) \sim t(N) \sim k(N)$	(increasing)	$i(N) = i_C(t)N^d$	(increasing)
w	productivity	w_D	(low)	$w(N)$ μ $i(N) \sim k(N)$	(increasing)	$w(N) = w_C(t)N^d$	(increasing)
t	time per function	t_D	(small)	$t(N) = A / d(N) \sim k(N)$	(increasing)	$t(N) = t_C(t)N^d$	(increasing)
c	cost per connection	c_D	(large)	$c(N) = Rj^2k(N)$	(increasing)	$c(N) = c_C(t)N^d$	(increasing)

Metcalfe's limit of the connected phase is obtained as $\delta \to 1$, see Figure 1. If information learned for the same functions across the network is redundant the total information in the connected phase is $I_C(N)=N\ i(N)d(N) \sim N$, which increases linearly with network size.

and the exponent (elasticity) δ is the rate of increase in connectivity with the size of the network and is assumed to be independent of N. This scale invariant form is predicted by urban scaling theory [41] and is observed in urban cell phone networks [42].

Then, I will assume that the number of functions accessible to each individual remains constant but becomes increasingly available to him/her through network connections [Fig. 1(b)], such that $k(N)d(N) = A$, which is independent of N. This is necessary because specializing individuals require access to functions they once maintained, e.g., a car mechanic needs access to a food producer and *vice versa*. This means specifically that the rate of increase in average individual specialization with system size equals in magnitude that of the increase in connectivity: $\dot{k}/k = -\dot{d}/d$, where the dots denote derivatives with respect to N. This quantitative behavior is observed, for example, in patterns of professional specialization and social connectivity in U.S. cities [17], and may already be present in simpler human societies [43]. Then, we conclude that $d(N) = A/k(N) = d_C N^{-\delta}$, so that each individual on average specializes in a smaller number of tasks. As a result he/she spends on average an amount of time on each task, $t(N) = T/d(N) = t_C N^{\delta}$, which increases proportionally to connectivity. Finally, we should expect that the total new information acquired (human capital) is proportional to the time on task, that is, $i(N) \propto t(N) = i_C N^{\delta}$, and that productivity is proportional to such information, $w(N) \propto i(N) = w_C N^{\delta}$, and thus, ultimately to connectivity.

These patterns are summarized in Fig. 1(c), as the rising green line, and in Table 1. They hold for any other dependence of connectivity on size, not just the illustrative power law, and express how a dynamical phase of network growth can take hold and lead to associated increases in overall information content, functional diversity, and individual productivity.

I can now also show how this reasoning maps to network effects and give an estimate of the value of connectivity. According to the so-called Metcalfe's law [44]: the *value* of connectivity is proportional to the square N^2 of the number of connected nodes N in a network. Up to a multiplicative constant, this is the maximum number of connections that can be realized in a network of N nodes. This result is readily obtained from the reasoning given above in the particular case when the parameter $\delta \to 1$, implying that each individual is connected to all others, and that, as a consequence, $k = N - 1$, and the total number of connections K is $K = Nk/2 \propto N(N-1)/2 \sim N^2$. I will refer to the regime $\delta \to 1$ as Metcalfe's limit; see Fig. 1(c).

Though Metcalfe's argument captures the potential maximum number of connections in a network, it is neither a direct measure of value nor a realistic assignment of the number of nodes that can actually be connected in a large system [41]. In order to derive the extent of connectivity in large networks, we must consider its costs, to which we now turn.

C. Costs of Connectivity

Connectivity is generally costly. A sense of the problem can be obtained by considering connectivity as a physical act of exchange. This exchange may involve the motion of physical goods, the transportation of people, or the transmission of information. Thus, each process of connectivity is mediated by a current j. In all macroscopic networks, there are dissipative energy losses associated with such exchanges that depend on the current as $C = Rj^2$. Here R is a resistance, set by whatever dissipative processes are relevant for the given exchange (e.g., friction in transportation or resistive losses for electricity) [41]. A well-known example is energy dissipation in electrical circuits (Joule's law). This reasoning also shows that energy costs are inevitable in irreversible exchanges (as a consequence of the second law of thermodynamics), whereas the translation of these costs into other units, such as money, may vary more widely, for example, as a result of the price of energy and choices of technology.

Thus, in our networked model, we expect a dissipative cost associated with each connection (as an independent current) proportional to the square of the intensity of the exchange times the relevant resistance parameter. The result in our network setting is that the cost per node $c(N) = Rj^2 k(N) = c_C N^\delta$. (This becomes $c(N) = c_C N$ in Metcalfe's limit.) Thus, the ratio c_C/k_C measures the average cost per connection: It is independent of the size of the system, but is in general a function of time through technological and organizational innovations.

D. Cost-Benefit Analysis and Network Transitions

Finally, we can assemble the general picture of benefits and costs of connectivity in informational networks to derive average expectations about when the connected network phase and its associated dynamics of learning and increasing productivity may take place.

First, let us consider the net gains $w_n(N)$ from connectivity as $w_n(N) = w(N) - c(N)$. Using the expressions derived above, I can write $w_n(N)$ as

$$w_n(N) = (w_c - c_c)N^\delta = \left(\frac{w_C}{k_C} - \frac{c_C}{k_C}\right)k(N).$$

From this expression, we immediately see that the connected network phase does not always pay off; see Fig. 1(c). In particular, if the costs exceed the productivity per connection, $c_C \geq w_C$, the behavior typical of the connected phase cannot develop at all. Only in the opposite regime, when connectivity becomes inexpensive in units of productivity, does this network become able to probe its dynamical learning regime and explore the advantages of the division and interdependence of labor and information. To see the nature of this transformation more clearly, I note that the transition should occur when net productivity in the connected phase can be larger than in the disconnected phase, or, mathematically

$$w_n(N) > w_D \rightarrow k(N) > \frac{w_D k_C}{w_C - c_C} \Leftrightarrow N > \left(\frac{w_D}{w_C - c_C}\right)^{\frac{1}{\delta}}$$

where I used the power-law parameterization of $k(N)$ in the last expression. Otherwise, a different finite threshold for N would result in similar qualitative behavior. This condition shows that, everything else being equal, the transition to the connected phase is inexorable as network size N increases. Although this phenomenon may be related to ideas of development through population pressure [45] and circumscription theory [46], clearly the key to these structural changes in complex networks can be more general and the underlying necessary conditions likely more subtle. However, this transition can also be produced at fixed size N as the tradeoff between advantages and costs of connectivity in the connected phase shift. Thus, this transition may be smooth or sudden (as in a tipping point) depending on whether the system is able to immediately capitalize on the new available dynamics of connectivity, or remains temporarily stuck in the disconnected phase even as favorable circumstances for the shift develop [47]. It should also be clear that, while the mean-field model introduced here allows us to anticipate a transition between the two network phases, it does not reveal its detailed nature. For example, we cannot tell whether this transition is smooth, or first or second order, in the language of physics. Such questions will require a full statistical approach to the two network phases.

The role of technology in complex informational networks now starts to come into focus: by creating a positive benefit-to-cost tradeoff for connectivity across the largest possible realm of interactions, technological change can place networked systems on a path of collective learning and of gains in terms of diversity and productivity. Technologies here should be understood in the broadest possible sense, from cultural and political institutions that help realize the benefits of social interdependence to fast computing, or large memory and bandwidth. Most often, transformative technologies must operate both in the purely technological realm and on extant social conditions.

III. EXAMPLES: CITIES, GLOBAL ONLINE NETWORKS, AND WIKIPEDIA

To anchor some of these concepts, I now discuss some of the network properties of a few important sociotechnical systems, from cities to online informational networks.

A. Cities

Cities are first and foremost social networks of people. Throughout most of human history humans lived in small, self-sufficient groups and assumed stereotypical roles within these groups, adapted to their natural environments, e.g., as hunter-gatherers or subsistence farmers [Fig. 1(a)] [20]. This small socioeconomic connectivity and lack of strong interdependence is a general characteristic of simple human societies, from those early in history to those that remain rural and "underdeveloped" today. Thus, the (more) disconnected state of human societies is visible both cross-sectionally from large cities to the smallest towns, across places characterized by very different levels of socioeconomic development and over time. It is an important question whether the change from this immemorial way of life to modern interconnected societies can be understood as a true network transition of the type introduced above. The apparent stability of simple subsistence societies suggests the hypothesis that maybe it is.

In contrast to disconnected networks, urbanizing societies are characterized by growing settlements where frequent social interactions with many different people become possible [48], [49]. Through co-location and faster internal transportation, cities reduce the cost of social

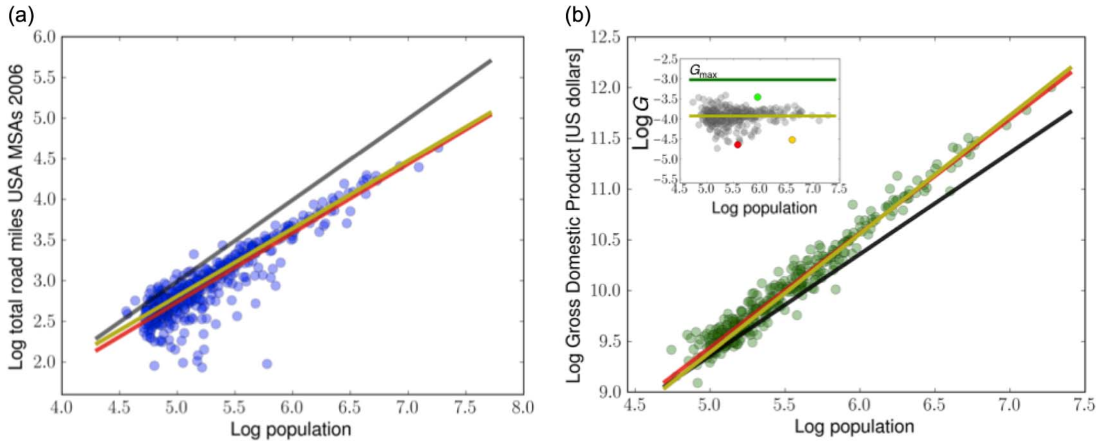


Fig. 2. ***General scaling properties of urban networks (adapted from [41]). (a) Total lane miles (volume) of roads in U.S. metropolitan areas (MSAs) in 2006 (blue dots). Data for 415 urban areas was obtained from the Office of Highway Policy Information from the Federal Highway Administration. Lines show a best fit to a power-law scaling relation with $b = 0.85$ (95% CI $= [0.81, 0.89]$, $R^2 = 0.65$). (b) Gross metropolitan product of MSAs in 2006 (green dots). Data obtained for 363 MSAs from U.S. Bureau of Economic Analysis. Lines describe best fit (red) to data, $b = 1.13$ (95% CI $= [1.10, 1.16]$, $R^2 = 0.96$). The black line shows a linear relation with unit slope. The yellow line shows the theory's prediction [41]. The inset shows the estimate of G for 313 U.S. MSAs, measured as the product of GDP and road volume, both per capita. Observed values of G for different cities are city size independent and cluster around a mean value expressing maximum net productivity, bounded by a maximum (green line) as predicted by the theory.***

connectivity and set in motion profound structural transformations in human societies [49], [50]. Thus, although the character of particular cities at specific times in history may appear more specific, urban centers are at heart the general means to the open-ended processes of human social development possible in the networked phase of social systems. In this light, as I emphasized elsewhere [41], [49], cities are the ultimate general-purpose "social reactors" [41], [49].

Because of the growing availability of data on urban quantities in cities all over the world, we are now able to measure quantitatively the expression of these processes and write them in terms of the framework developed in Section II. To set the stage, consider the results shown in Fig. 2, for the scaling of total road surface and gross domestic product (GDP) of U.S. metropolitan statistical areas (a definition of cities as interacting networks); see [41] and [50]. Both these quantities show average scaling behavior according to an expression of the form

$$Y(N,t) = Y_0(t)N^b$$

with $b = 1 + \delta$, where δ is the exponent introduced in the previous section. This follows from the fact that per capita quantities (denoted by lower case letters) and total quantities are proportional up to a factor of N, e.g., $Y = Ny$. Y, expressed as GDP, is a measure of (economic) output, and its main component, wages, which measures labor productivity, scales in the same way [41]. The plots of Fig. 2 show that the value of the exponent $\delta \simeq 1/6$, as predicted by theory [41], is based on the integrated modeling of cities as social networks embedded in infrastructural space and subject to both connectivity gains and costs as a result of social interactions in space.

In this framework, the exponents and prefactors can be computed explicitly from models of the general geometry of urban infrastructure; see [41] for details. Briefly, in this context, the exponent δ measures the rate of densification of people in public spaces (proxied by the infrastructure network volume V_n), $n = N/V_n = n_0 N^\delta$, where n is density. V_n, shown in Fig. 2(a), is in turn the result of building an infrastructure network with the general properties that 1) it connects all spatial parts of the city together; and 2) remains open-ended in the sense that it can be expanded gradually as the city grows. While some aspects of this calculation are specific to cities as spatial structured systems, I want to emphasize that it is the increase in density in some space that promotes greater average connectivity per unit time, $k(N) \sim n(N) \sim N^\delta$, as we observe directly in cities via cell phone connectivity [42]. I show in Sections II-B and II-C that this effect persists in online networks, with different exponent values, where no explicit reference to physical space is necessary.

Second, the value of prefactors such as $Y_0(t)$ is the result of the optimization of the cost benefit structure of gains minus costs following from social connectivity across all dimensions of life, that is, of the factor $w_C - c_C$. In cities, it can be shown that both benefits and costs of interaction depend on certain parameters associated with mobility (of people, goods, and information), and that these in turn are a function of the adaptation of human behavior to the characteristics of infrastructure and *vice versa* [41]. As a result it can be shown that the analog quantity to $w_C - c_C$, G in Fig. 2(b) (inset), is city size independent and has a well-defined maximum value across cities. Different cities manifest a slightly different value of their G, around this value (dots in the inset). In an analogous way, we should expect that different instantiations of networks in the connected phase should manifest values of their $w_C - c_C$ around a mean value.

Finally, while it remains difficult to find direct proxies for learning and the specialization of knowledge in cities, one can measure $d(N)$ through the consideration of the statistics of employment across a large number of different of professional occupations [17]. This analysis, applied to U.S. metropolitan areas, confirms the expectation for $d(N) \sim N^{\delta}$, with $\delta \approx 1/6$, and of labor productivity (measured through wages) exhibiting the property that $w(N)d(N) = A$, a constant independent of N, and thus that $w(N) \sim N^{\delta}$, as hypothesized above. Similarly, the social connectivity of urban social networks can be estimated at the individual level using cell phone networks [42], resulting on the scaling behavior of connectivity consistent with these observations, as predicted by theory [41].

The consistency of these ideas must continue to be explored empirically, especially in different nations and through more microscopic studies. These results do, however, provide an important illustration of the dynamics of the connected network phase proposed above and of its development over space and time in circumstances that are fundamental to understanding human sustainable development, technological change, and economic growth.

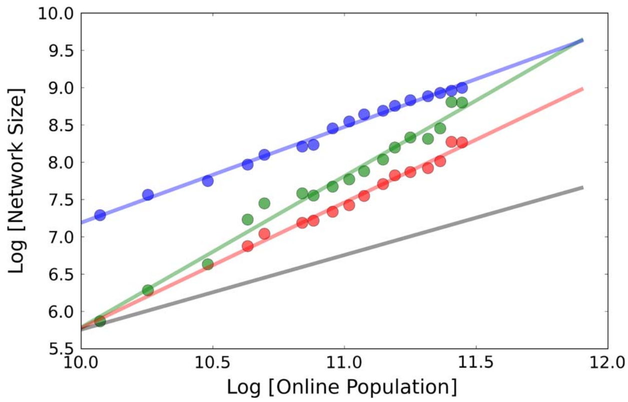


Fig. 3. ***Scaling of global computer networks with online population size. The size of the Internet, measured in terms of DNS hosts (blue) is characterized by an exponent 1.28 (95% CI = [1.22, 1.34], $R^2 = 0.99$), while the growth of the WWW, in terms of an estimate of total webpages (green), is characterized by an exponent 2.03 (95% CI = [1.88, 2.17], $R^2 = 0.98$) and of active pages (red) by an exponent 1.68 (95% CI = [1.55, 1.82] $R^2 = 0.98$). In all cases, the size of online networks has been growing superlinearly with the number of Internet users, indicating that more pages and more computation is effectively used per capita as the network grows, much like in other open-ended social systems (e.g., cities). Exponents are manifestly different from those observed for cities.***

B. The Internet and the World Wide Web

Over the last 20 years, progress in computing and telecommunication technologies has enabled unprecedented growth in connectivity between distant people. These technologies are also creating networks of knowledge that are, in specific senses, external to individual humans and their social networks and where information is instead encoded in webs of interlinked "documents," without explicit spatial location. The Internet and the World Wide Web (WWW) embody these global changes and continue to evolve from more specific and smaller networks to new and more pervasive realms.

It is, therefore, interesting to study the evolution of these networks in light of the general concepts developed above: To what extent are the Internet and the WWW examples of the connected phase dynamics of informational networks? What sort of productivity and learning are they creating?

Such enterprise is fraught with conceptual caveats and empirical limitations, some of them discussed in here below and in Section VI. The main difference from the analysis of urban data is that we do not have the ability to perform cross-sectional analysis with the Internet or the WWW, and must hope instead that their time evolution can give us a sense of the size dependence of their informational properties. If, in addition to growth in their size, the prefactors change exponentially in time, however, they will contribute to the magnitude of the exponents and are likely to result in their overestimation.

Perhaps surprisingly, the actual size and connectivity of these networks remain largely unknown both because they have become immensely large, but also because of their decentralized and bottom-up dynamics (see data sources in Section VI). Mapping them requires in practice that the entire network is visited, node by node, in order to estimate their global structure. Nevertheless, several surveys give us a sense at least of the broad structural dynamics of these networks as their user-base has increased over time. Fig. 3 shows the number of DNS hosts and two estimates of the number of webpages as functions of the total worldwide online population, in analogy to Fig. 2 for cities. The first interesting feature of Fig. 3 is that, in both cases, we observe clear superlinear scaling $(b > 1, \delta > 0)$. While this sort of behavior for DNS hosts suggests an increase in task load (cost) on servers with each additional person online, the growth in number of pages is especially interesting as it suggests a (more strongly superlinear) increase in content, and thus perhaps in the productivity of the system. The number of total webpages (which we should think functionally as links between users), in particular, exhibits scaling with an exponent consistent with Metcalfe's law. However, this webpage count is plagued by certain spurious effects related to incentives to artificially create pages (see Section VI) so that the number of active sites, which scale with a smaller superlinear exponent, may be a more accurate measure.

In any case, it is interesting that we infer from these results that the number of webpages has grown with online population size at a rate much faster than social connections with the population size of urban agglomerations. In this sense, each individual may typically have access to more pieces of information online and be able to specialize on his/her own production to a larger extent.

To my knowledge, this is the first demonstration of pervasive superlinear scaling of the Internet and the WWW with online population size. However, these measures remain very rough estimates of the growth of these networks, and it would be very interesting to revisit the present results with better data. In addition, it would be desirable to obtain other measures more directly related to online connectivity, information, and individual attention and of their evolution over time.

C. Wikipedia

Another, more particular online network example, where more thorough measures of network properties are available, is Wikipedia, the online encyclopedia. Wikipedia started in January 2001 and has grown spectacularly even since, comprising currently of over 30 million articles across its large set of different languages.

Wikipedia is not a general-purpose network aimed at increasing general productivity or connectivity. Its goal is to create encyclopedia articles collaboratively, through the contributions of anyone who wishes to participate. In this sense, nodes, treated as articles, do increase their information content over time through the intervention (edits) of human contributors. Thus, even though nodes do not learn *per se*, we can treat them in analogy to the scheme developed above, with humans being a part of the connectivity structure (and bearing some of the costs) of creating and growing the network and its nodes.

The growth of the body of cross-referenced articles hence created does then provide us with a picture of how information as a whole increases and how its productivity in terms of impact may change in tandem. This happens in two ways: 1) through the iterative process of improvement of each document (which is a process of collective learning encoded as the article); and 2) through the linkages (connectivity) that an entry establishes to others, both internal and external to Wikipedia. Thus, it is this network of documents that encodes information and it is its change that represents the process of learning. Although readers of Wikipedia may also benefit (and learn) from this encoding of knowledge, contributors to Wikipedia may not individually possess all the knowledge that a single page reflects (that is the point of the collaborative model). This turns the process of learning in cities (and the parallel suggestive structure of scaling in the WWW) upside down and suggests that the best measure for the size of Wikipedia are articles and that their connectivity is supplied by human contributors as well as document links, not *vice versa*.

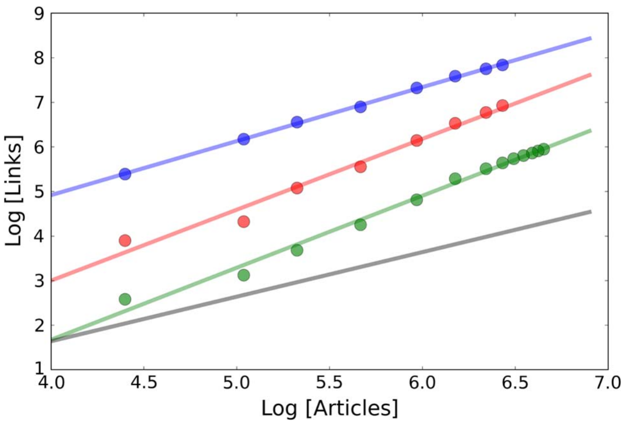


Fig. 4. ***Contributors and external and internal links to Wikipedia articles scale superlinearly. The number of contributors (green) scales with the number of Wikipedia articles with exponent $b = 1.61$ (95% CI $= [1.51, 1.72]$, $R^2 = 0.99$). The number of external links (red) scales approximately in the same way with exponent $b = 1.59$ (95% CI $= [1.40, 1.79]$, $R^2 = 0.98$). Finally, the larger number of internal links (blue) scales more slowly with an exponent $b = 1.21$ (95% CI $= [1.18, 1.24]$, $R^2 = 0.99$). If we interpret these quantities as different measures of connectivity between articles, we see that they all scale with exponents larger than those observed for social connectivity in cities.***

If we adopt this perspective, we find scaling results that broadly agree with those we invoked for cities and the WWW, but with different exponents. Fig. 4 shows how the network connectivity, measured in terms of human contributors as well as internal and external links, grows superlinearly with the number of articles. Fig. 5 helps justify the identification of Wikipedia contributors with page links, by showing that these scale linearly (proportionally) to each other. Fig. 5 also shows that the number of edits in Wikipedia is proportional to the number of contributors, supporting the assertion that the cost of connectivity per link is constant in N, as is human effort in cities [41]. Finally, Fig. 6 shows that a proxy for the productivity of an average article (and of each contributor) increases superlinearly, at least in terms of audience reach. This also establishes that the benefit of creating a connection (the effort of a contributor) is outpaced by its benefits in terms of audience reach, suggesting indeed that Wikipedia is an informational network expressing the connected phase dynamics.

These examples supply evidence that informational networks typically enable payoffs that are superlinear on the number of learner elements and that they are limited in their growth primarily by the cost of establishing and maintaining this connectivity. Whenever benefits outstrip costs, these networks can connect (and in some cases grow) explosively [Fig. 1(c)]. Eventually, they may equilibrate to a scale invariant regime where costs and benefits scale in the same (superlinear) way.

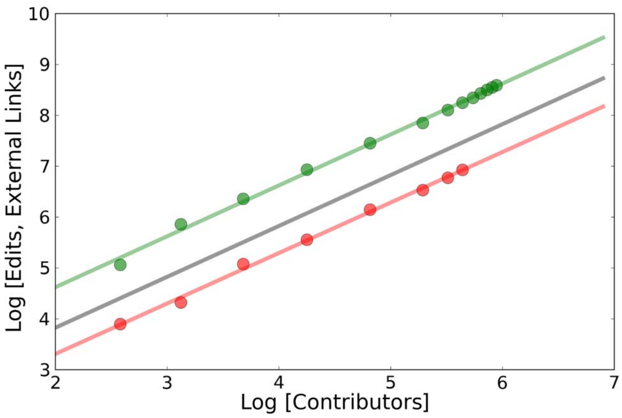


Fig. 5. ***Total number of edits (green) and external links (red) to Wikipedia articles is proportional to the number of individual contributors. The green line shows the best fit to the number of edits versus contributors with exponent $b = 1.00$ (95% CI = $[0.97, 1.03]$, $R^2 = 0.99$). The red line shows the best fit to the number of external links versus contributors with exponent $b = 0.99$ (95% CI = $[0.95, 1.03]$, $R^2 = 0.99$). The black line shows exact proportionality $b = 1$, for comparison.***

IV. TECHNOLOGICAL TRENDS IN COMPLEX INFORMATIONAL NETWORKS

Finally, I would like to discuss several scenarios for the interplay between physical and informational networks and the impact of changing technology on their evolution.

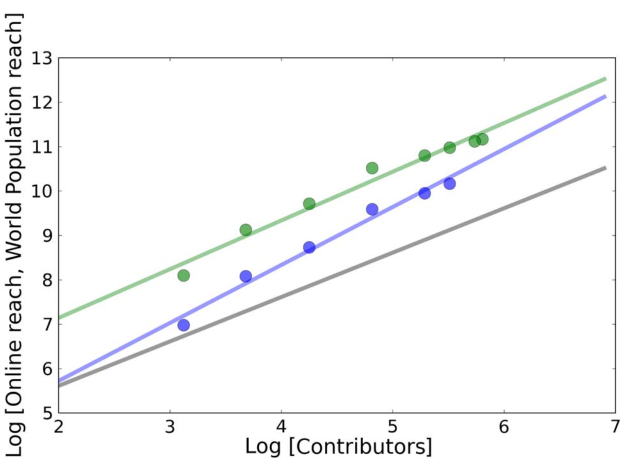


Fig. 6. ***Audience reach of contributors to Wikipedia increases superlinearly. The green line shows the best fit to Alexa online reach surveys, which estimates the number of Internet users who read Wikipedia, with $b = 1.10$ (95% CI = $[0.95, 1.25]$, $R^2 = 0.97$). The blue line shows the best fit to the reach in terms of total worldwide online population with $b = 1.31$ (95% CI = $[1.23, 1.38]$, $R^2 = 0.98$). The blue line is steeper than the green line because it accounts for the growth of the online population versus the total world population, which in 2013 is estimated at 39%.***

A. The Death of Distance? Can Cyberspace Replace Physical Space?

I argued above that the dynamics characteristic of the connected network phase are primary to its "means" or processes. This means that whether this dynamics can be realized in cities, where space and infrastructure play an essential role, or online, where attention and time are apparently more relevant "spaces," is a secondary consideration.

Except, of course, to the extent that one of these modes overcomes disadvantages of another and may hence substitute it altogether. In this light, a common question is whether the Internet and information and telecommunication technologies (ICTs) can eliminate the need for cities [51], or, instead, whether these very different networks play synergetic and mutually supportive roles. The hypothesis that physical proximity becomes unnecessary is often described by the concept of the "death of distance" [51]–[56].

Research over the last decade has pursued answers to these questions with mixed results. Two general findings seem to stand out and are worth noting: 1) online and ICT networks are local; and 2) the uses of the Internet and of local social networks tend to be integrated and tend to complement, rather than substitute one another. On the first point, it has been found empirically that more online content is available in larger cities [52], [53], so that these new technologies tend to reinforce rather than replace the connectivity dynamics of larger places. This means in particular that maps, services, etc., are disproportionally available online if they stem from larger cities. In this vein, it is probably interesting to remark that previous introductions of informational and telecommunications technologies, from the newspaper and the postal service to the telegraph and the telephone, were always skewed toward larger environments, and not simply for economic reasons related to their cost [25]. On the second point, most findings are both intuitive and obvious: whether for shopping and commerce [53]–[55] or for telecommuting [56], new ICTs are extremely useful in helping organize the complex life patterns typical of larger cities, including the fine temporal coordination involved in meetings. In this sense, new informational technologies are most useful in the most intensely connected network phases, which typically are to be found in large cities [41], precisely because they lower the cost of their pervasive connectivity.

Despite these findings, the question remains whether vastly improved telecommunication technologies, capable of reproducing the nuances of sharing space and meeting face to face, can one day replace personal travel [56]. While there is no reason to exclude such possibility, my guess is that all connected networks will tend to mesh together and reinforce one another and that substitution is only possible when new modes fully include and transcend the advantages of previous modes of interaction and learning.

B. Technological Change, Connectivity, and Information in Constant-Size Networks

In the framework of Section II, I entangled the issue of increases in the information content and productivity of a networked information system together with changes in its size N. Here, I briefly discuss how these processes can take place independently, that is, even when N is constant. The key issue is the variation in the remaining parameters, prefactors, and exponents. This can generally be captured by their time dependence.

Consider then changes in the baseline productivity per social connection w_C and in the costs of connectivity c_C. The first can be affected by the adoption or improvement of new production methods or increased learning rates, while the latter may change due to improvements in transportation or through the development of social and political institutions that handle conflict resolution more effectively [41]. Thus, the crucial consideration is how baseline net productivity $w_C - c_C$ changes over time.

An increase in this quantity moves the green curve in Fig. 1(c) up, shifting its intersection with the horizontal line to the left and consequently makes the connected network phase more advantageous sooner; that is, at smaller network sizes N. Conversely, a decrease in the baseline net productivity moves the green line down and reduces the attractiveness of connectivity, delaying its onset to larger N.

Thus, we see that certain temporal shifts in these baseline parameters are sufficient, at fixed N, to produce the transition to large-scale learning and growth in complex networks. This is analogous to ideas of intensive economic growth [12] through technological change in theories of endogenous economic growth [14]. Endogenizing growth, in turn, would require additional models for how N changes over time and its associated costs, and whether a partial allocation of productivity to such costs can generate a virtuous cycle of growth and learning, or will instead fizzle out. Such elaborations are left to future work.

C. Resilience of Connected Informational Networks

Finally, it is important to discuss some of the pitfalls of the simple dynamics of differentiation, learning, and growth described above. A more complete consideration of other important factors, not treated in detail in this paper, is addressed in Section V.

First, the path of increasing individual specialization may, in some circumstances, lead to static arrangements that stall processes of open-ended learning and productivity increases. In real circumstances in human societies, extreme labor and knowledge specialization are sometimes only possible inside vertically integrated organizations (hierarchies), such as those of large firms and of universities and research laboratories [57]. Such environments can promote the stability and continuity necessary for the pursuit of more speculative R&D or for extreme specialization, say in an assembly line, in ways that economic markets often do not support. The danger of this internal specialization is that knowledge hence created remains tied to its very specific context and cannot be used in new generative ways in large networks [58]. In science and technology, specific communication channels, such as the publication of scientific manuscripts and patents, help bridge this gap, but much knowledge still remains tacit and local. This difficulty may prove more severe in large-scale manufacturing where factory-floor workers are typically at once very specifically matched to their tasks and redundant with each other and with automated solutions. In these circumstances, labor is not free to specialize further, or to learn in ways that may benefit the individual over the long term or the networked system as a whole.

This creates an apparent contradiction: While the creation and full use of specialized knowledge often requires protective environments inside stable organizations, its value depends on broad openness and exchange at the network level. These two processes, taken together, suggest that dynamics of formation and dissolution of organizations (such as hierarchies) are likely necessary for new information to be created and for it to realize its full value. This concept is partly captured by the idea of quasi-decomposability of hierarchies [59]. This can be achieved through open and dynamical labor markets, entrepreneurship, and processes by which knowledge can be accumulated in stable but open ways, for example, through open platforms, such as in online Wikis and open-source repositories.

A second issue relates to the resilience of the connected network phase. It should be clear that the disconnected phase, though it is characterized by low productivity and information content, is generally very robust to the loss of nodes. This is a direct result of its informational redundancy at the node level, a mechanism that is often employed in engineering solutions to ensure against random local failures [60].

The source of resilience of the connected state emerges not so much from its structure, but rather from its dynamics: In this phase, the loss of nodes implies some loss of information, and the loss of connections may reverse the process of learning; but these processes are, to some extent, reversible. The idea is that being ultimately dynamical the system can adjust to a loss in system size by tracing its evolution backward. This implies some loss in knowledge as well as some degradation of productivity, but still maintains the system ready to bounce back and reevolve again.

Thus, the question of resilience in informational networks is whether, upon a shock, the connected system can degrade gracefully and bounce back quickly. Anecdotal evidence from recent disasters in cities suggests that people can take up many of the functions that are usually performed by infrastructure and services [61]. Examples are walking or bicycling as a substitute to mass transit. But, the possibility remains that fast and reversible adaptation may not always be possible and that sudden, hard to reverse transitions may occur, accompanied by the destruction of network connectivity and of critical information. The exploration of such important phenomena requires further development of the ideas discussed here in terms of their statistical dynamics.

V. DISCUSSION AND CONCLUSION

In this paper, I proposed general and intentionally very simple dynamics of network development and showed that under certain circumstances nodes capable of learning may remain disconnected, redundant, and relatively unproductive, while in other situations, they may develop pervasive connectivity and embark on an open-ended trajectory of growth and development.

These ideas were largely conceived in the context of cities, though I argued here that they likely reflect more fundamental concepts common across several disciplines and apply also, at a different scope and speed, to worldwide computer networks and webs of information, such as the Internet and Wikipedia.

While I hope that the framework developed here provides a simple integrated perspective on the open-ended evolution of complex informational networks, much remains to be done to further develop these ideas theoretically, to test them empirically, and to make them useful in guiding policy and technological development.

Regarding theory, a development of more formal definitions of information and learning processes and their relationship to network structure is clearly necessary. This issue is already reasonably well developed technically on some fronts, such as in information theory [40], [62], Bayesian networks [63], and related artificial neural networks [33]. Nevertheless, the concept of node differentiation and learning through changes in their internal states occurring in tandem with related changes in network connectivity may provide new formal developments in such models, which remain to date incapable of the kind of learning that is common for humans and their social networks [33].

On this vein, the consideration of network structures that go beyond changes in average connectivity will almost certainly be necessary. Connectivity and learning costs, as well as productivity increases, may be facilitated by the creation and dissolution of formal organizations, such as firms, civic associations, and other local and online communities. Such entities are characterized by complex internal organization, often in the form of emergent hierarchical structures. Consequently, I expect that, besides average increases in connectivity, local network structures should play an essential role in the development of any real system. It may be interesting to develop a version of the quantities considered here that is local within larger networks (such as firms inside markets) and thus gain access to the more explicit dynamics of spreading processes that may result from successful local adaptations as they gain a foothold across larger networks.

Finally, I purposefully avoided explicit considerations of the cost and benefits of learning and information. This is an old problem, bypassed in the original formulation of information theory [40], [62] but essential to the understanding of evolution and development in complex systems. Such costs and benefits can be measured in many different units, such as money or energy. However, it seems obvious that such measures fail to capture much of the dynamics of online networks. In these cases, different measures of value may be necessary, such as, for example, attention (i.e., human time) [64], but it remains unclear whether they are sufficient.

A crucial issue going forward is our ability to understand and predict quantitatively how these connected informational networks will evolve, including the detailed consequences of specific new technologies. At present, in society and in technology, we are dependent on human creativity and intervention to achieve most processes of learning and knowledge recombination. Understanding these processes in more integrated and detailed ways will ultimately create a new science of complexity that can explain massive evolving informational networks such as ecosystems, cities, and the World Wide Web. It will also create a world of more intelligent technologies, able to coevolve and coadapt in connection with humans and with other natural and artificial networked complex systems.

VI. MATERIALS AND METHODS

Data on the size of the WWW (the number of total websites and the number of active websites) shown in Fig. 2 were obtained through the Web Server Survey by Netcraft (news.netcraft.com). The methods for this survey are described at www.netcraft.com/active-sites, including the distinction between all sites and active sites. The need for this distinction, which is aimed at excluding some websites created automatically and others targeted at increasing search engine visibility, was less crucial in the early years of the WWW. The size of the Internet, measured by the number of domain name system (DNS) hosts, was obtained from the Internet domain survey by the Internet Systems Consortium (ftp.isc.org/www/survey/reports/2013/07). The number of worldwide people online was obtained from estimates in the International Telecommunications Union annual reports (www.itu.int) in percent and converted into a number using estimates of the world's total population obtained from geohive (www.geohive.com). I emphasize that the nature of most measurements of the size of the WWW, the Internet, and online population is obtained through surveys and may be subject to incompleteness and biases. Wikipedia size and usage data were obtained from online Wikipedia statistics (stats.wikimedia.org). ■

Acknowledgment

The author would like to thank J. Lobo and S. Ortman for discussions on theories of economic growth and innovation.

REFERENCES

[1] M. Castells, *Rise of the Network Society*. New York, NY, USA: Wiley-Blackwell, 1996.

[2] M. Hilbert and P. López, "The world's technological capacity to store, communicate, and compute information," *Science*, vol. 332, pp. 60–65, Apr. 2011.

[3] B. Goldstein, Ed., *Beyond Transparency: Open Data and the Future of Civic Innovation*. San Francisco, CA, USA: Code for America Press, 2013.

[4] J. Bollen *et al.*, "Clickstream data yields high-resolution maps of science," *PLoS ONE*, vol. 4, Mar. 2009, e4803.

[5] W. B. Arthur, *The Nature of Technology: What It Is and How It Evolves*. New York, NY, USA: Free Press, 2011.

[6] J. Holston, *Insurgent Citizenship*. Princeton, NJ, USA: Princeton Univ. Press, 2008.

[7] W. R. Easterley, *The Elusive Quest for Growth*. Cambridge, MA, USA: MIT Press, 2002.

[8] J. Höller *et al.*, *From Machine-to-Machine to the Internet of Things: Introduction to a New Age of Intelligence*. Amsterdam, The Netherlands: Elsevier, 2014.

[9] National Research Council, Policy Division, Board on Sustainable Development, *Our Common Journey: A Transition toward Sustainability*. Washington, DC: National Academy Press, 1999.

[10] M. Newman, *Networks: An Introduction*. Oxford, U.K.: Oxford Univ. Press, 2010.

[11] C. Hidalgo, B. Klinger, A.-L. Barabasi, and R. Hausmann, "The product space conditions the development of nations," *Science*, vol. 317, pp. 482–487, Jul. 2007.

[12] R. J. Barro and X. I. Sala-i-Martin, *Economic Growth*. Cambridge, MA, USA: MIT Press, 2003.

[13] S. A. Frank, "The Price equation, Fisher's fundamental theorem, kin selection, and causal analysis," *Evolution*, vol. 51, pp. 1712–1729, Dec. 1997.

[14] P. M. Romer, "The origins of endogenous growth," *J. Econ. Perspect.*, vol. 8, pp. 3–22, Dec. 1994.

[15] F. Hayek, "The use of knowledge in society," *Amer. Econ. Rev.*, vol. 35, pp. 519–530, Sep. 1945.

[16] G. S. Becker and K. M. Murphy, "The division of labor, coordination costs, and knowledge," *Quart. J. Econ.*, vol. CVII, pp. 1137–1160, Nov. 1992.

[17] L. M. A. Bettencourt, H. Samaniego, and H. Youn, "Professional diversity and the productivity of cities," *Sci. Rep.*, vol. 4, p. 5393, Jun. 2014.

[18] E. L. Glaeser and M. G. Resseger, "The complementarity between cities and skills," *J. Reg. Sci.*, vol. 50, pp. 221–244, Jan. 2010.

[19] S. E. Page, *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools and Societies*. Princeton, NJ, USA: Princeton Univ. Press, 2007.

[20] A. W. Johnson and T. Earle, *The Evolution of Human Societies*. Stanford, CA, USA: Stanford Univ. Press, 2000.

[21] L. Mumford, *The City in History: Its Origins, Its Transformations, and Its Prospects*. New York, NY, USA: Harcourt, Brace & World, 1961.

[22] F. Braudel, *Civilization & Capitalism 15th–18th Century, Vol. 1, 2, 3*. New York, NY, USA: Harper, 1979.

[23] P. W. Anderson, "More is different," *Science*, vol. 177, pp. 393–396, Aug. 1972.

[24] R. Florida, *Cities and the Creative Class*. New York, NY, USA: Routledge, 2004.

[25] P. Hall, *Cities in Civilization*. New York, NY, USA: Pantheon Books, 1998.

[26] J. Jacobs, *The Economy of Cities*. New York, NY, USA: Vintage, 1970.

[27] D. Tapscott and A. D. Williams, *Wikinomics: How Mass Collaboration Changes Everything*. New York, NY, USA: Portfolio Trade, 2010.

[28] R. Botsman and R. Roo, *What's Mine Is Yours: The Rise of Collaborative Consumption*. New York, NY, USA: Harper Business, 2010.

[29] W. B. Arthur, *Increasing Returns and Path Dependence in the Economy*. Ann Arbor, MI, USA: Univ. Michigan Press, 1994.

[30] L. Sveikauskas, "The productivity of cities," *Quart. J. Econ.*, vol. 89, pp. 393–413, Aug. 1975.

[31] S. S. Rosenthal and W. C. Strange, "The determinants of agglomeration," *J. Urb. Econ.*, vol. 50, pp. 191–229, Sep. 2001.

[32] G. Duranton and D. Puga, "Diversity and specialisation in cities: Why, where and when does it matter?" *Urban Studies*, vol. 37, pp. 533–555, Aug. 2000.

[33] S. Russel and P. Norvig, *Artificial Intelligence: A Modern Approach*. New York, NY, USA: Prentice-Hall, 2009.

[34] A. Smith, *An Inquiry Into the Nature and Causes of the Wealth of Nations*. London, U.K.: W. Strahan, 1776.

[35] E. Durkheim, *The Division of Labor in Society*. New York, NY, USA: Free Press.

[36] M. Weber, *The Theory of Social and Economic Organization*, T. Parsons, Ed. New York, NY, USA: Free Press, 1947.

[37] K. Arrow, "The economic implications of learning by doing," *Rev. Econ. Studies*, vol. 29, pp. 155–173, Jun. 1962.

[38] B. Nagy, J. D. Farmer, Q. M. Bui, and J. E. Trancik, "Statistical basis for predicting technological progress," *PLoS ONE*, vol. 8, Feb. 2013, e52669.

[39] K. A. Ericsson, N. Charness, P. J. Feltovich, and R. R. Hoffman, Eds., *The Cambridge Handbook of Expertise and Expert Performance*. Cambridge, U.K.: Cambridge Univ. Press, 2006.

[40] D. MacKay, *Information Theory, Inference, and Learning Algorithms*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[41] L. M. A. Bettencourt, "The origin of scaling in cities," *Science*, vol. 340, pp. 1438–1441, Jun. 2013.

[42] M. Schläpfer *et al.*, "The scaling of human interactions with city size," *J. Roy. Soc. I.*, vol. 6, Sep. 2014, 20130789.

[43] M. A. Kline and R. Boyd, "Population size predicts technological complexity in Oceania," *Proc. Roy. Soc. Biol. Sci.*, vol. 277, pp. 2559–2564, Apr. 2010.

[44] Wikipedia, "Metcalfe's Law." [Online]. Available: http://en.wikipedia.org/wiki/Metcalfe's_law

[45] E. Boserup, *Population and Technological Change: A Study of Long-Term Trends*. Chicago, IL, USA: Univ. Chicago Press, 1981.

[46] R. L. Carneiro, "The transition from quantity to quality: A neglected causal mechanism in accounting for social evolution," *Proc. Nat. Acad. Sci. USA*, vol. 97, pp. 12 926–12 931, Nov. 2000.

[47] D. Achlioptas, R. D'Souza, and J. Spencer, "Explosive percolation in random networks," *Science*, vol. 323, pp. 1453–1455, Mar. 2009.

[48] L. Wirth, "Urbanism as a way of life," *Amer. J. Soc.*, vol. 44, pp. 1–24, Jul. 1938.

[49] L. M. A. Bettencourt, "The kind of problem a city is," SFI working paper 2013-03-008, 2013. [Online]. Available: http://www.santafe.edu/media/workingpapers/13-03-008.pdf

[50] L. M. A. Bettencourt, J. Lobo, D. Helbing, C. Kühnert, and G. B. West, "Growth, innovation, scaling, and the pace of life in cities," *Proc. Nat. Acad. Sci. USA*, vol. 104, pp. 7301–7306, Apr. 2007.

[51] S. Kirsch, "The incredible shrinking world? Technology and the production of space," *Environ. Plann. D*, vol. 13, pp. 529–555, Mar. 1995.

[52] J. Kolko, "The death of cities? The death of distance? Evidence from the geography of commercial Internet usage," in *The Cities in the Global Information Society: An International Perspective*. Cambridge, MA, USA: MIT Press, 1999, pp. 73–98.

[53] T. Sinai and J. Waldfogel, "Geography and the Internet: Is the Internet a substitute or a complement for cities?" *J. Urban Econ.*, vol. 56, pp. 1–24, Jul. 2004.

[54] E. Tranos and P. Nijkamp, "The death of distance revisited: Cyber-place, physical and relational proximities," *J. Reg. Sci.*, 2013, DOI: 10.1111/jors.12021.

[55] J. Pons-Novell and E. Viladecans-Marsal, "Cities and the Internet: The end of distance?" *J. Urban Tech.*, vol. 13, pp. 109–132, Dec. 2010.

[56] P. Zhu, "Are telecommuting and personal travel complements or substitutes?" *Ann. Reg. Sci.*, vol. 48, pp. 619–639, Apr. 2012.

[57] K. Arrow, *The Limits of Organizations*. New York, NY, USA: Norton, 1974.

[58] J. Jacobs, *The Nature of Economies*. New York, NY, USA: Random House, 2000.

[59] H. A. Simon, "The architecture of complexity," in *The Sciences of the Artificial*. Cambridge, MA, USA: MIT Press, 1969.

[60] S. Naeem, "Species redundancy and ecosystem reliability," *Conserv. Biol.*, vol. 12, pp. 39–45, Feb. 1998.

[61] E. G. T. Tompson, J. Benz, J. Agiesta, K. Cagney, and M. Meit, "Resilience in the wake of Superstorm Sandy," The Associated Press/NORC Center for Public Affairs Research. [Online]. Available: http://www.apnorc.org/projects/Pages/resilience-in-the-wake-of-superstorm-sandy.aspx

[62] E. T. Jaines, *Probability Theory: The Logic of Science*. Cambridge, U.K.: Cambridge Univ. Press, 2003.

[63] J. Pearl, *Causality: Models, Reasoning, and Inference*. Cambridge, U.K.: Cambridge Univ. Press, 2000.

[64] F. Wu and B. A. Huberman, "Novelty and collective attention," *Proc. Nat. Acad. Sci. USA*, vol. 104, pp. 17599–17601, Oct. 2007.

ABOUT THE AUTHOR

Luís M. A. Bettencourt was born in Lisbon, Portugal, in 1969. He received the Licenciatura degree (five years) in engineering physics from the Instituto Superior Técnico, University of Lisbon, Lisbon, Portugal, in 1992 and the Ph.D. degree in theoretical physics from Imperial College, University of London, London, U.K., in 1996.

From 1997 to 2003, he was a Postdoctoral Fellow at The Institute for Theoretical Physics, Heidelberg, Germany; a Director's Fellow at Los Alamos National Laboratory, Los Alamos, NM; and a Postdoctoral Associate at the Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA, USA. In 2003, he became a Staff Member in the Computer and Computational Sciences at the Theoretical Division, Los Alamos National Laboratory. Since November 2011, he has been a Professor at the Santa Fe Institute, Santa Fe, NM, USA. Currently, he works on complex systems theory and different aspects of cities and urbanization. His research emphasizes the creation of new interdisciplinary synthesis to describe cities in quantitative and predictive ways, informed by the growing availability of empirical data worldwide. He is particularly interested in the interplay between information, structure, and scale in setting the properties of diverse complex systems.

INVITED PAPER

Overlapping Communities Explain Core–Periphery Organization of Networks

The paper presents a new paradigm for uncovering the modular structure of complex networks.

By Jaewon Yang and Jure Leskovec

ABSTRACT | Networks provide a powerful way to study complex systems of interacting objects. Detecting network communities—groups of objects that often correspond to functional modules—is crucial to understanding social, technological, and biological systems. Revealing communities allows for analysis of system properties that are invisible when considering only individual objects or the entire system, such as the identification of module boundaries and relationships or the classification of objects according to their functional roles. However, in networks where objects can simultaneously belong to multiple modules at once, the decomposition of a network into overlapping communities remains a challenge. Here we present a new paradigm for uncovering the modular structure of complex networks, based on a decomposition of a network into any combination of overlapping, nonoverlapping, and hierarchically organized communities. We demonstrate on a diverse set of networks coming from a wide range of domains that our approach leads to more accurate communities and improved identification of community boundaries. We also unify two fundamental organizing principles of complex networks: the modularity of communities and the commonly observed core–periphery structure. We show that dense network cores form as an intersection of many overlapping communities. We discover that communities in social, information, and food web networks have a single central dominant core while communities in protein–protein interaction (PPI) as well as product copurchasing networks have small overlaps and form many local cores.

KEYWORDS | Community detection; core–periphery structure; ground-truth communities; networks

I. INTRODUCTION

Networks provide a way to represent systems of interacting objects where nodes denote objects (people, proteins, webpages) and edges between the objects denote interactions (friendships, physical interactions, links). Nodes in networks organize into communities [1], which often correspond to groups of nodes that share a common property, role or function, such as functionally related proteins [2], social communities [3], or topically related webpages [4]. Communities in networks often overlap as nodes might belong to multiple communities at once. Identifying such overlapping communities in networks is a crucial step in studying the structure and dynamics of social, technological, and biological systems [2]–[5]. For example, community detection allows us to gain insights into metabolic and protein–protein interactions (PPIs), ecological foodwebs, social networks like Facebook, collaboration networks, information networks of interlinked documents, and even networks of copurchased products [6]–[12]. In particular, communities allow for analysis of system properties that cannot be studied when considering only individual objects or the entire system, such as the identification of module boundaries and relationships and the classification of objects according to their functional roles [13]–[17].

Manuscript received March 20, 2014; revised August 1, 2014 and October 15, 2014; accepted October 15, 2014. Date of publication November 11, 2014; date of current version November 18, 2014. This work was supported by the National Science Foundation (NSF) under Career Award IIS-1149837, by Defense Advanced Research Projects Agency (DARPA) XDATA, by GRAPHS, by the Alfred P. Sloan Fellowship, and by the Microsoft Faculty Fellowship.

J. Yang is with the Department of Electrical Engineering, Stanford University, Stanford, CA 94305 USA (e-mail: jayang@cs.stanford.edu).
J. Leskovec is with the Department of Computer Science, Stanford University, Stanford, CA 94305 USA (e-mail: jure@cs.stanford.edu).

Digital Object Identifier: 10.1109/JPROC.2014.2364018

Here we explore the community structure of a number of networks from many domains. We distinguish between structural and functional definitions of communities [18]. Communities are often structurally defined as sets of nodes with many connections among the members of the set and few connections to the rest of the network [1]. Communities can also be defined functionally based on the function or role of community's members. For example, functional communities may correspond to social groups in social networks, scientific disciplines or research groups in scientific collaboration networks, and biological modules in PPI networks. The basic premise of community detection is that these functional communities share some common structural signature, which allows us to extract them from the network structure.

Based on this distinction one can state that the goal of community detection is to build a bridge between network structure and function, that is, to identify communities based on the network structure with the aim that such structurally identified communities would correspond to functional communities.

In this paper, we build on the above view of network community detection and develop community detection methods that identify functional communities based on their structural connectivity patterns. We identify networks where we can obtain reliable external labels of functional communities. We refer to such explicitly labeled functional communities as ground-truth communities [18]. We study structural properties of ground-truth functional communities and find that they exhibit a particular structural pattern. We discover that the probability of nodes being connected increases with the number of ground-truth communities they share. Our observation means that nodes residing in overlaps of ground-truth communities are more densely connected than nodes in the nonoverlapping parts of communities. Interestingly, we also find that assumptions behind many existing overlapping community detection methods lead to the opposite conclusion that the more communities a pair of nodes shares, the less likely they are to be connected [6]–[11]. Thus, as a consequence, many overlapping community detection methods may not be able to properly detect ground-truth communities.

Based on the above observations, we develop a new overlapping community detection method called the community-affiliation graph model (AGM), which views communities as overlapping "tiles" and the tile density corresponds to edge density [19]. Fig. 1 illustrates the concept. Our methodology decomposes the network into a combination of overlapping, nonoverlapping, and hierarchically organized communities. We compare AGM to a number of widely used overlapping and nonoverlapping community detection methods [6], [7], [10], [20] and show that AGM leads to more accurate functional communities. On average, AGM gives 50% relative improvement over existing methods in assigning nodes to their ground-truth communities in social, coauthorship, product copurchasing, and biological networks.

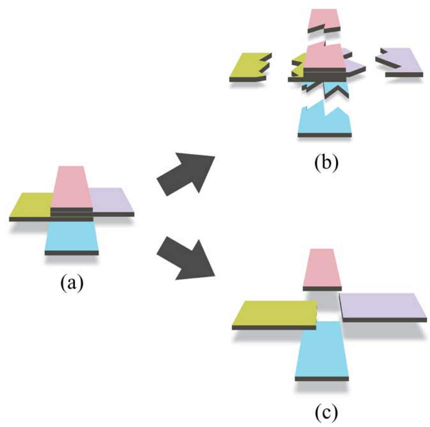


Fig. 1. ***Communities as tiles. (a) Communities in networks behave as overlapping tiles. (b) Many methods view communities as clusters with a homogeneous edge density and thus they may break the tiles. (c) Our AGM methodology successfully decomposes the network into different tiles (communities).***

Finally, we unify two fundamental organizing principles of complex networks: overlapping communities and the commonly observed core–periphery structure. While network communities are often thought of as densely linked clusters of nodes, in core–periphery network structure, the network is composed of a densely connected core and a sparsely connected periphery [21]–[23]. Many large networks may exhibit core–periphery structure. The network core was traditionally viewed as a single giant community, and, therefore, it was conjectured that the core lacks internal communities [24]–[27]. We unify those two organizing principles and show that dense network cores form as a result of many overlapping communities. Moreover, we find that foodweb, social, and web networks exhibit a single dominant core while PPI and product copurchasing networks contain many local cores formed around the central core.

Our methodology to decompose networks into communities provides a powerful tool for studying social, technological, and biological systems by uncovering their modular structure. Our work represents a new way of studying networks of complex systems by bringing a shift in perspective from defining communities as densely connected nodes to conceptualizing them as overlapping tiles.

II. FROM STRUCTURAL TO FUNCTIONAL DEFINITIONS OF COMMUNITIES

The traditional structural view of network communities is based on two fundamental social network processes: triadic closure [28] and the strength of weak ties theory [29], [30]. Under this view, structural communities are often defined as corresponding to sets of nodes with many

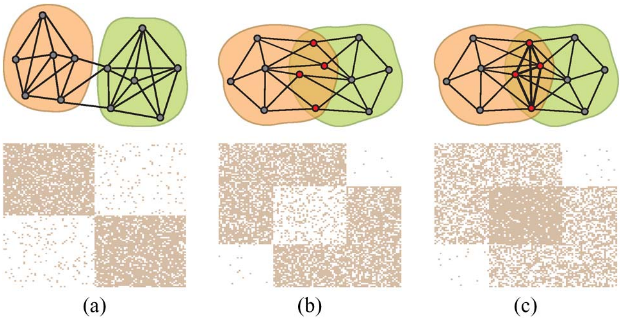


Fig. 2. ***Three structural definitions of network communities. Networks (top) and corresponding adjacency matrices (bottom), where rows/columns denote nodes and dots denote edges: (a) two nonoverlapping communities; (b) two overlapping communities where the overlap is less connected than the nonoverlapping parts of communities; and (c) two overlapping communities where the nodes in the overlap are better connected. Based on (c), we structurally define communities as analogous to "tiles," where community overlaps lead to higher density of edges.***

"strong" connections between the members of the community and few "weak" connections with the rest of the network [Fig. 2(a)]. However, in many domains, nodes may belong to multiple communities at once, and thus the notion of structural communities has also been extended to include overlapping, hierarchical, and disassortative community structures [6], [31]–[34].

Despite great progress in the field, we find that extending the traditional structural view to overlapping communities leads to an unnoticed consequence that nodes in community overlaps are less densely connected than nodes in the nonoverlapping parts of communities [Fig. 2(b)]. (Refer to the extended version of the paper [35] for details.) We find this hidden consequence to be present in many existing approaches to overlapping community detection [6]–[11].

We examine a diverse set of six networks drawn from a wide range of domains, including social, collaboration, and copurchasing networks for which we obtain explicitly labeled functional communities, which we refer to as the ground-truth communities [18]. For example, in social networks, we take ground-truth communities to be social interest-based groups to which people explicitly join, and in product networks, ground-truth communities are defined to correspond to categories of products [35]. Note we define ground-truth communities based on common functions or roles around which networks organize into communities [18]. Ground-truth communities are not defined based on some observed attribute or property that the nodes share (for example, age, gender, or hometown in a case of a social network) [6]. The idea behind ground-truth communities is that they would correspond to true functional modules in complex networks. While the obtained ground-truth labels may sometimes be noisy or incomplete, consistency and robustness of the results suggest that the ground-truth labels are overall reliable.[1]

By studying the structure of ground-truth communities we find that two nodes are more likely to be connected if they have multiple ground-truth communities in common (Fig. 3). For example, in the LiveJournal online social network (Table 1), the edge probability jumps from $\sim 10^{-6}$ for nodes that share no ground-truth communities to 0.1 for nodes that have one ground-truth community in common and keeps increasing all the way to 0.7 as nodes share more communities [Fig. 3(a)]. This implies that the area of overlap between two communities has a higher average density of edges than an area that falls in just a single community [Fig. 2(c)].

Our observation is intuitive and consistent across several domains. For example, proteins belonging to multiple common functional modules are more likely to interact [2], people who share multiple interests have a higher chance of becoming friends [36], and researchers with many common interests are more likely to collaborate [36].

A. Defining Structural Communities as Tiles

We think of communities as analogous to overlapping "tiles." Thus, just as the overlap of two tiles leads to a higher tile height in the overlapping area, the overlap of two communities leads to higher density of edges in the overlap. (Fig. 1 illustrates the concept.) The composition of many overlapping communities then gives rise to the global structure of the network.

Conceptually, our methodology represents a shift in perspective from structurally modeling communities as sets of densely linked nodes to modeling communities as overlapping tiles where the network emerges as a result of the overlap of many communities. Our structural definition of communities departs from the strength of weak ties theory [30] and is consistent with the earlier social network theory called the web of group affiliations [37], which postulates that edges arise due to shared community affiliations.

Our findings here also have implications for the understanding of homophily, which is one of the primary forces that shape the formation of social networks [36]. Homophily is the tendency of individuals to connect to others with similar tastes and preferences. Based on [30], it has been commonly assumed that homophily operates in "pockets," and, thus, nodes that have neighbors in other communities are less likely to share the attributes of those neighbors [as in Fig. 2(a) and (b)]. In contrast, our results are implying pluralistic homophily where the similarity of nodes is proportional to the number of shared memberships/functions, not just their similarity along a

[1]Networks with ground-truth communities can be downloaded from http://snap.stanford.edu/agm.

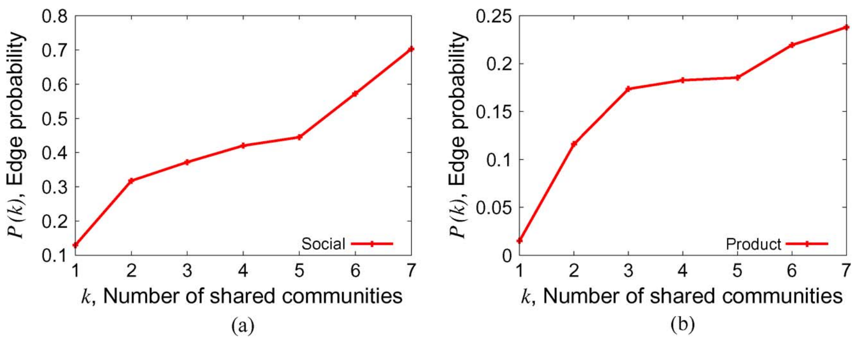


Fig. 3. ***Community overlaps have higher edge density than the nonoverlapping parts of communities. Edge probability $P(k)$ as a function of the number of common community memberships k (a) in the social network and (b) in the product copurchasing network (Table 1). Results in (a) and (b) suggest that, as nodes share multiple communities, they are more likely to be connected, which leads to higher edge density in community overlaps, as illustrated in Fig. 2(c).***

single dimension. In a multidimensional network, the most central nodes are those that have the most shared dimensions.

III. DECOMPOSITION OF NETWORKS INTO COMMUNITIES

In order to model communities in a network, we define the AGM [19]. In our model, edges of the underlying network arise due to shared community memberships [Fig. 4(a)] [38], [39]. The AGM parameterizes each community A with a single parameter p_A. Two nodes that belong to community A then form an edge in the underlying network with probability p_A. Each community A generates edges between its members independently; however, if two nodes have already been connected, then the duplicate edge is not included in the network.

The AGM naturally models communities with dense overlaps [Fig. 4(a) and (b)]. Pairs of nodes that belong to multiple common communities become connected in the underlying network with a higher probability, since for each shared community the nodes are given an independent chance of forming an edge.

The flexible nature of the AGM allows for modeling a wide range of network community structures, such as nonoverlapping, hierarchically nested, and overlapping communities [Fig. 4(c)–(e)]. Given a bipartite community affiliation graph and a probability p_A for each community A, the AGM allows us to generate synthetic networks with realistic community structures, a procedure useful in and of itself.

Using the AGM, we can also identify and analyze community structure of real-world networks. We accomplish decomposition of a given network into communities by fitting the AGM to the network with tools of statistical inference. We combine a maximum-likelihood approach with convex optimization and a Monte Carlo sampling algorithm on the space of community affiliation graphs [19], [35], [40]. This technique allows us to efficiently search for the community affiliation graph that gives the observed network the greatest likelihood. To automatically determine the number of communities in a given network, we

Table 1 Network Statistics and Properties of Detected Communities. We Consider the Facebook Ego-Network of a Particular User, the Full LiveJournal Online Social Network, the Florida Bay Foodweb Network, the Stanford University Web Graph, the Literature-Curated *Saccharomyces Cerevisiae* PPI Network, and the Amazon Product Copurchasing Network. Network Statistics: N: Number of Nodes; E: Number of Edges; $\langle C \rangle$: Average Clustering Coefficient; D: Effective Diameter; $\langle k \rangle$: Average Node Degree. Properties of Detected Communities: K: Number of Communities; $\langle S \rangle$: Average Detected Community Size; $\langle A \rangle$: Average Number of Community Memberships Per Node. The Networks Vary From Those With Modular to Highly Overlapping Community Structure and Represent a Wide Range of Edge Densities. While the Number of Communities Detected by AGM Varies, the Average Community Size Is Quite Stable Across the Networks. Average Number of Community Memberships Per Node Reveals That Communities in the Foodweb Overlap Most Pervasively, While in PPI and Social Networks Overlaps Are Smallest

Network	Properties of networks					Properties of detected communities		
	N	E	$\langle C \rangle$	D	$\langle k \rangle$	K	$\langle S \rangle$	$\langle A \rangle$
Facebook	183	2,873	0.56	2.80	31.40	4	70.8	1.5
Social network	3,997,962	34,681,189	0.28	6.47	17.35	29,774	83.3	0.6
Foodweb	128	2,075	0.33	1.90	32.42	5	54.4	2.1
Web graph	255,265	1,941,926	0.62	9.36	15.21	5,000	83.3	1.6
PPI network	1,213	2,556	0.33	10.50	4.21	40	31.6	1.0
Product network	334,863	925,872	0.40	15.00	5.53	9,020	50.0	1.3

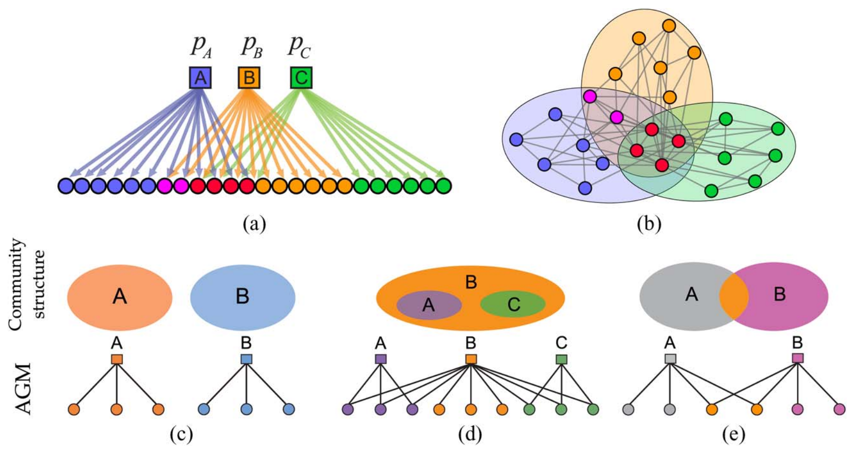


Fig. 4. ***Community-affiliation graph model (AGM) [19]. (a) Squares represent communities and circles represent the nodes of a network. Edges represent node community memberships. For each community A that two nodes share they create a link independently with probability p_A. The probability that a pair of nodes u, v creates a link is thus $p(u,v) = 1 - \prod_{A \in C_{uv}} (1 - p_A)$, where C_{uv} is the set of communities that u and v share. If u and v do not share any communities, we assume they link with a small probability ε. (b) Network generated by the AGM in (a). As pairs of nodes that share multiple communities get multiple chances to create edges, the AGM naturally generates networks where nodes in the community overlaps are more densely connected than the nodes in nonoverlapping regions. (c)–(e) AGM is capable of modeling any combination of (c) nonoverlapping, (d) hierarchically nested, as well as (e) overlapping communities.***

apply techniques from statistical regularization and sparse model estimation [35].

IV. ACCURACY OF DETECTED COMMUNITIES

Next, we aim to infer functional communities based only on the structure of a given unlabeled undirected network.

A. Qualitative Evaluation

As an illustrative example, we consider a Facebook friendship network of a single user's friends [Fig. 5(a) and Table 1]. In order to obtain labels for ground-truth communities, we asked the user to manually organize his Facebook friends into communities. The user classified his friends into four communities corresponding to his

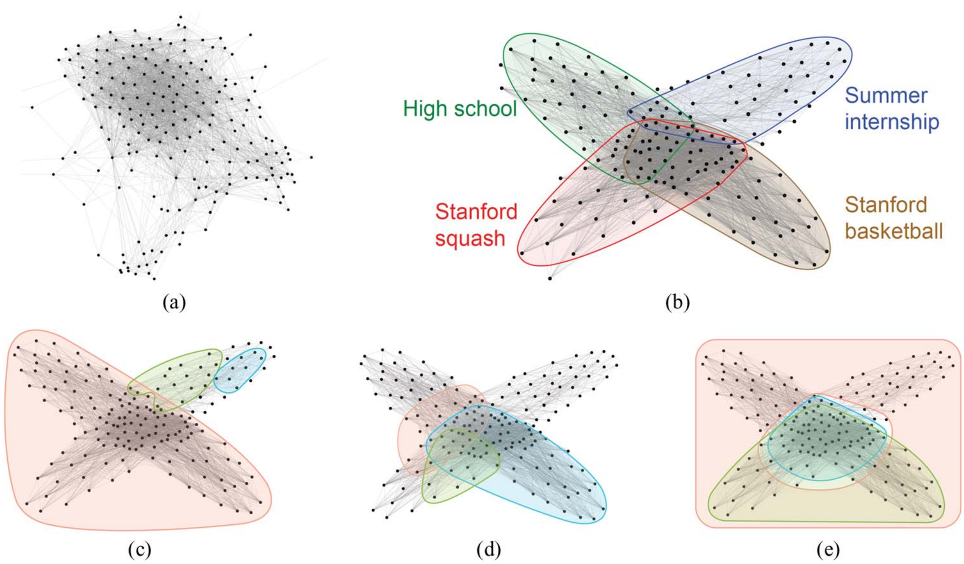


Fig. 5. ***An example on a Facebook friendship network of a particular user. (a) Facebook friendship network of a single user. (b) The same network but with communities explicitly labeled by the user: high school friends, colleagues at the workplace, and university friends with whom the user plays basketball and squash. Communities are denoted by filled regions. Note that nodes in the overlap of communities have higher density of edges. (c)–(e) Results of applying (c) clique percolation, (d) link clustering, and (e) mixed-membership stochastic block model to the Facebook network.***

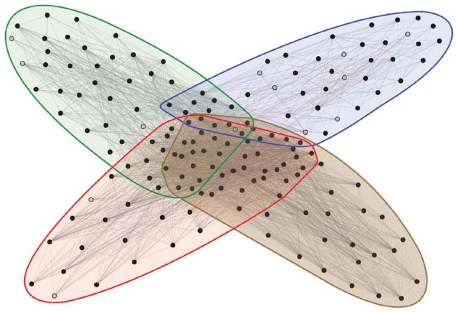


Fig. 6. ***AGM on the Facebook network from Fig. 5. AGM successfully decomposes the network into different tiles (communities) and correctly determines community overlaps as well as community memberships for 94% of the nodes.***

high school, workplace, and two communities of university friends. The visualization of the same network using communities in Fig. 5(b) shows that the network in Fig. 5(a) is in fact composed of the overlaps of the four communities. In this example, the goal of community detection is to identify the communities in Fig. 5(b) based only on the connectivity structure of the network in Fig. 5(a).

Due to an implicit assumption that nodes in community overlaps are less densely connected than nodes in the nonoverlapping parts of communities [Fig. 2(a)], many overlapping community detection approaches [6]–[11] fail to properly detect communities in this network. For example, Fig. 5(c)–(e) illustrates the result of applying clique percolation [10], link clustering [6], and the mixed-membership stochastic block (MMSB) model [7] to the Facebook network in Fig. 5(a). We also give a formal argument that explains the behavior of these methods in Appendix I-A and the extended version [35].

However, when we use the AGM to analyze the Facebook network, the AGM automatically detects four communities (Fig. 6), which is the same as the number identified by the user. Moreover, the communities detected by the AGM nearly perfectly correspond to communities identified by the user. The AGM correctly determines community overlaps and community memberships for 94% of the user's friends.

B. Quantitative Evaluation

We also perform a large-scale quantitative evaluation on AGM on biological, social, collaboration, and product networks where functional communities are explicitly labeled [18]. The networks represent a wide range of sizes and edge densities, as well as amounts of community overlap. We compare the AGM to a number of widely used overlapping and nonoverlapping community detection methods [6], [7], [10], [20] and quantify the correspondence between the explicitly labeled ground-truth communities and the communities detected by a given method. The performance metrics quantify the accuracy of the method in assigning nodes to their ground-truth communities. (Refer to Appendix I-B for further details.)

On a set of social, collaboration, and product networks, AGM on average outperforms existing methods by 50% in four different metrics that quantify the accuracy in assigning nodes to their ground-truth communities [Fig. 11(a)]. In particular, AGM gives a 50% relative improvement over clique percolation [10]. Link clustering [6] detects overlapping as well as hierarchical communities and AGM improves 61% over it. Similar levels of improvement are achieved when comparing AGM to other overlapping and nonoverlapping methods [7], [20]. Furthermore, AGM gives a 14% relative improvement over link clustering using the same networks and the same data-driven benchmarks as used in the link clustering work [6].

Furthermore, we also experiment with AGM on a set of four different biological PPI networks. Remarkably, even though AGM was developed based on insights gained on primarily social networks, we find that AGM performs surprisingly well on biological networks as well. As performance metrics, we compute the average statistical significance of detected communities (p-value) for the three types of gene ontology (GO) [41]. We consider negative logarithm of average p-values for each of the three GO term types as three separate scores. On average, the AGM outperforms link clustering by 150%, the clique percolation method (CPM) by 163%, Infomap by 148%, and the MMSB model for 12 times [Fig. 11(b)]. Further experimental details are in the Appendixes III, IV, and [35].

Overall, the AGM approach yields substantially more accurate communities. The success of our approach relies on the AGM's flexible nature, which allows the AGM to decompose a given network into a combination of overlapping, nonoverlapping, and hierarchical communities.

V. COMMUNITIES, PLURALISTIC HOMOPHILY, AND CORE–PERIPHERY STRUCTURE

The AGM also makes it possible to gain well-founded insights into the community structure of networks. In particular, we discover that overlapping communities lead to a global core–periphery network structure. Core–periphery structure captures the notion that many networks decompose into a densely connected core and a sparsely connected periphery [21], [22]. The core–periphery structure is a pervasive and crucial characteristic of large networks [23], [24], [42].

We discover that a network core forms as a result of pluralistic homophily where the connectedness of nodes is

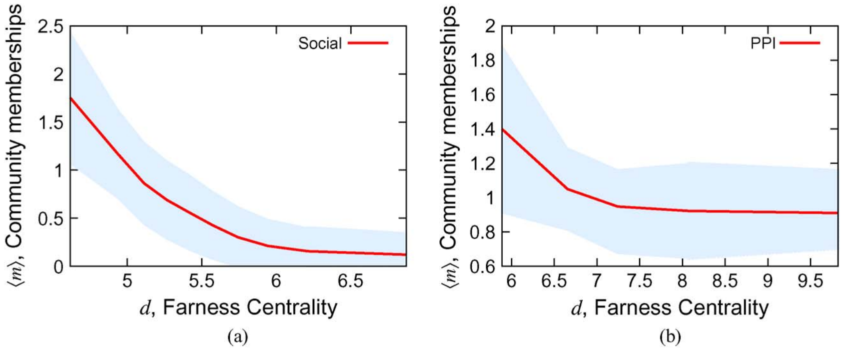


Fig. 7. ***Overlapping communities lead to global core–periphery network structure. The average (and the tenth percentiles) of the number of community memberships $\langle m\rangle(d)$ as a function of its farness centrality d, defined as the average shortest path length of a given node to all other nodes of the network [3]. (a) LiveJournal social network. (b) Saccharomyces cerevisiae PPI network. The number of community memberships of a node decreases with its farness centrality. Nodes that reside in the center of the network (and have small shortest path distances to other nodes of the network) belong to the highest number of communities. This means that core–periphery structure forms due to community overlaps. Communities in the periphery tend to be nonoverlapping while communities in the core overlap pervasively.***

proportional to the number of shared community memberships, and not just their similarity along a single dimension or community. Thus, the network core forms as a result of many overlapping communities. The average number of community memberships of a node decreases with its distance from the center of the network (Fig. 7). Moreover, the edge likelihood increases as a function of community memberships (Fig. 3). Thus, the nodes in the center of the network have higher density of edges than nodes in the periphery. Therefore, we show that even in the presence of many communities, pluralistic homophily leads to dense community overlaps, which cause a global core–periphery network structure.

A further examination of the amount of community overlap reveals that social, web, and foodweb networks in Table 1 have a single central dominant core [Fig. 8(a)]. On the other hand, communities in protein and product networks have small overlaps and also form many local cores [Fig. 8(b)]. In particular, protein communities only slightly overlap and form local cores as well as a small global core [Fig. 8(d)]. Small overlaps of protein communities can be explained by the fact that communities act as functional modules, and it would be hard for the cell to independently control heavily overlapping modules [2], [6]. Communities of copurchased products can also be thought of as functional modules since the products in a community are bought together for a specific purpose. On the other hand, foodweb communities overlap pervasively while forming a single dominant core. This leads to a flowerlike overlapping community structure [Fig. 8(c)], where tiles (communities) overlap to form a single core of the network. The heavily overlapping foodweb communities form due to the closed nature of the studied Florida Bay ecosystem [43]. Web communities overlap moderately and form a single global core. Many of these communities form around common interests or topics, which may overlap with each other [4].

VI. CONCLUSION

We note that our approach builds on the previous work on community detection [6]–[16]. We examine an implicit assumption of sparsely connected community overlaps and find that regions of the network where communities overlap have higher density of edges than the nonoverlapping regions.

We then rethink classical structural definitions of communities and develop the AGM, which models structural communities as overlapping tiles. Using our well-founded approach, we find that all networks considered in this study exhibit a core–periphery structure where nodes that belong to multiple communities reside in the core of the network. However, networks have different kinds of core–periphery structure depending on the mechanism for community formation in the networks. Dense community overlaps also explain the mixed success of present community detection methods when applied to large networks [24], [27].

Our work also enhances our understanding of homophily as one of the most fundamental social forces. Homophily in networks has been traditionally thought to operate in small pockets/clusters. Thus, nodes that have neighbors in other communities were considered less likely to share properties of those neighbors. In contrast, our results are implying pluralistic homophily where the similarity of nodes' properties is proportional to the number of shared community memberships. In a network, the most central nodes are those that have the most shared properties/functions/communities with others. More generally, our work provides a shift in perspective from conceptualizing communities as densely connected sets of nodes to

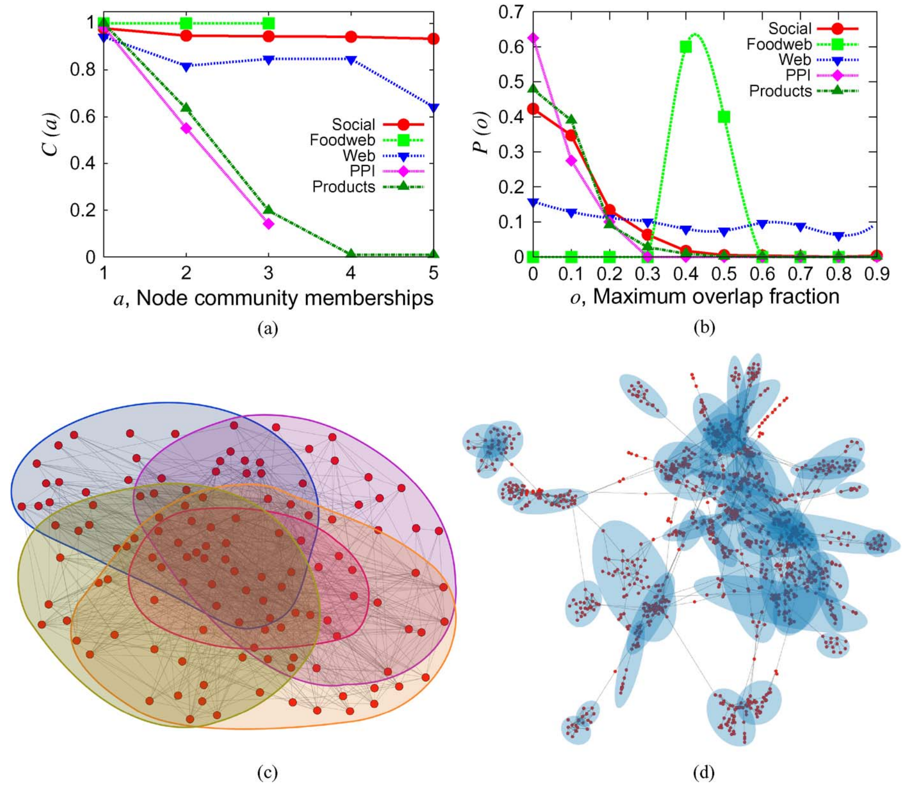


Fig. 8. *Primary and secondary cores in networks. (a) The fraction of nodes $C(a)$ in the largest connected component of the induced subgraph on the nodes who belong to at least a communities. By thinking of a network as a valley where peaks correspond to cores and peripheries to lowlands, our methodology is analogous to flooding lowlands and measuring the fraction of the largest island. A high $C(a)$ means that there is a single dominant core (peak), while a low $C(a)$ suggests the existence of nontrivial secondary cores. (b) Probability density $P(o)$ of the maximum overlap o. Maximum overlap o_A of a given community A is defined as the fraction of A's nodes that are in the overlap with any other community. Communities in the PPI, social, and product copurchasing networks are mostly nonoverlapping whereas the communities in the foodweb and the web graph are pervasively overlapping. (c) Communities detected by the AGM in the foodweb form a single central core. (d) Communities in the PPI network form many secondary cores.*

defining them as overlapping tiles and represents a new way of studying complex systems. ■

APPENDIX I
DETECTING DENSELY OVERLAPPING COMMUNITIES

Here, we show that three popular community detection methods, clique percolation (CPM) [10], [44], link clustering (LC) [6], and stochastic block model [7], [45], cannot properly detect communities with dense overlaps.

A. Clique Percolation

First, we analyze the CPM and show that it may not properly detect two overlapping communities from Fig. 2(c). The CPM has a single input parameter k which determines the size of the maximal cliques that the algorithm looks for. For example, Fig. 9 shows the result of CPM on the network of Fig. 2(c) where the overlap between the two communities is denser than the individual communities. When $k = 3$, the CPM finds a community that covers the whole network because the clique in the overlap connects the cliques in the left community and the right community, whereas the CPM finds a community of the overlap when $k = 4$.

In addition to CPM, there are many other overlapping community detection methods that are based on expanding the maximal cliques. These methods (for example, Greedy

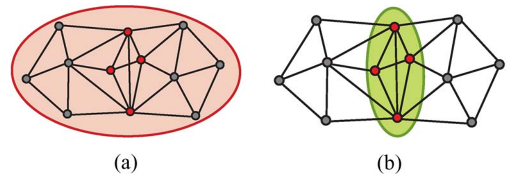


Fig. 9. *Clique percolation method cannot detect communities with dense overlaps. Given a network with two communities and a dense overlap, clique percolation method would report a community that (depending on the parameter settings) either (a) includes both communities, or (b) it would find a small community consisting only of the overlap: (a) $k_{CPM} = 3$; (b) $k_{CPM} = 4$.*

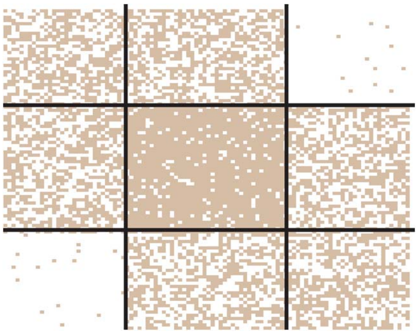


Fig. 10. *The result of the stochastic block model and the mixed-membership stochastic block model on a network of two communities with dense overlap. The adjacency matrix of the network in Fig. 2(c) is shown, and the bold lines denote the three partitions discovered by the stochastic block models, where the overlap is confused as a separate community.*

clique expansion [46] and EAGLE [47]) also suffer from the same problem.

B. Stochastic Block Models

We show that three variants of stochastic block models are unable to correctly discover communities with dense overlaps: the traditional stochastic block model [45], the degree-corrected stochastic block model [48], and the MMSB model [7]. Based on the input matrix from Fig. 2(c), all three models identify three blocks, as illustrated in Fig. 10. The reason for this is that the edge probability between two nodes that belong to communities A and B is weighted average of $P(A,A)$ and $P(B,B)$, where $P(X,Y)$ is an edge probability between a node in community X and a node in community Y. This means that the edge probability between the two nodes that share multiple communities is smaller than the maximum of $P(A,A)$ and $P(B,B)$ (due to the weighted summation). Therefore, the edge probability between overlapping nodes cannot be higher than the edge probability between nodes in an individual community. We also note that in principle one could apply postprocessing of communities detected by stochastic block models to identify which of the detected structural communities actually correspond to overlaps of functional communities. However, it is not immediately clear how to develop such postprocessing method.

C. Link Clustering

Last, we show that link clustering [6] is not able to correctly detect overlapping communities with dense overlaps. Link clustering performs hierarchical clustering with the edges of the given network based on the Jaccard similarity between the adjacent nodes of the edges. Since edge density in the area of community overlap is higher, this means that the Jaccard similarity between the adjacent nodes will be higher, which in turn means that link clustering will identify the edges in the overlap as a separate community. (Refer to the extended version [35] for a more formal argument.)

APPENDIX II
METRICS OF COMMUNITY DETECTION ACCURACY

We focus the evaluation of community detection methods on their ability to correctly identify overlapping ground-truth communities.

To quantify the performance, we measure the level of agreement between the detected and ground-truth communities. Given a network $G(V,E)$, we consider a set of ground-truth communities C^* and a set of detected communities $\hat{C}$, where each ground-truth community $C_i \in C^*$ and each detected community $\hat{C}_i \in \hat{C}$ is defined by a set of its member nodes. To compare $\hat{C}$ and C^*, we use four performance metrics.

1) Average F1 score [49]: We compute $F_g(C_i) = \max_j F1(C_i, \hat{C}_j)$ for each ground-truth community C_i and $F_d(\hat{C}_i) = \max_j F1(C_j, \hat{C}_i)$ for each detected community $\hat{C}_i$, where $F1(S_1, S_2)$ is the harmonic mean of precision and recall between two node sets S_1, S_2. The average F1 score is $(1/2)(\bar{F}_g + \bar{F}_d)$ where $\bar{F}_g = (1/|C^*|)\sum_i F_g(C_i)$ and $\bar{F}_d = (1/|\hat{C}|)\times \sum_i F_d(\hat{C}_i)$.
2) Omega index [50]: For each pair of nodes $u, v \in V$, we define C_{uv} to be the set of ground-truth communities to which both u and v belong and $\hat{C}_{uv}$ to be the set of detected communities to which both nodes belong. Then, the omega index is $(1/|V|^2)\sum_{u,v\in V} \mathbf{1}\{|C_{uv}| = |\hat{C}_{uv}|\}$.
3) Normalized mutual information [12]: We compute $1-(1/2)(H(C^*|\hat{C})+H(\hat{C}|C^*))$, where $H(A|B)$ is the extension of entropy when A, B are sets of sets [12].
4) Accuracy in the number of communities: $1-(||C^*|-|\hat{C}||)/|C^*|$, which is the relative error in predicting the number of communities.

APPENDIX III
APPLYING AGM TO SOCIAL, PRODUCT, AND COLLABORATION NETWORKS

Fig. 11(a) displays the composite performance of each of the five methods over the six networks with ground-truth communities. Overall, we note that AGM gives best overall performance on all networks, except the Amazon, where it ties with MMSB. Furthermore, AGM detects highest quality communities for most individual performance metrics in all networks. On average, the composite performance of AGM is 3.40, which is 61% higher than that of link clustering (2.10), 50% higher than that of CPM (2.41), 30% higher than that of Infomap, and 8% higher than that of MMSB (3.25). The absolute average value of omega index of AGM over the six networks is 0.46, which is 21%

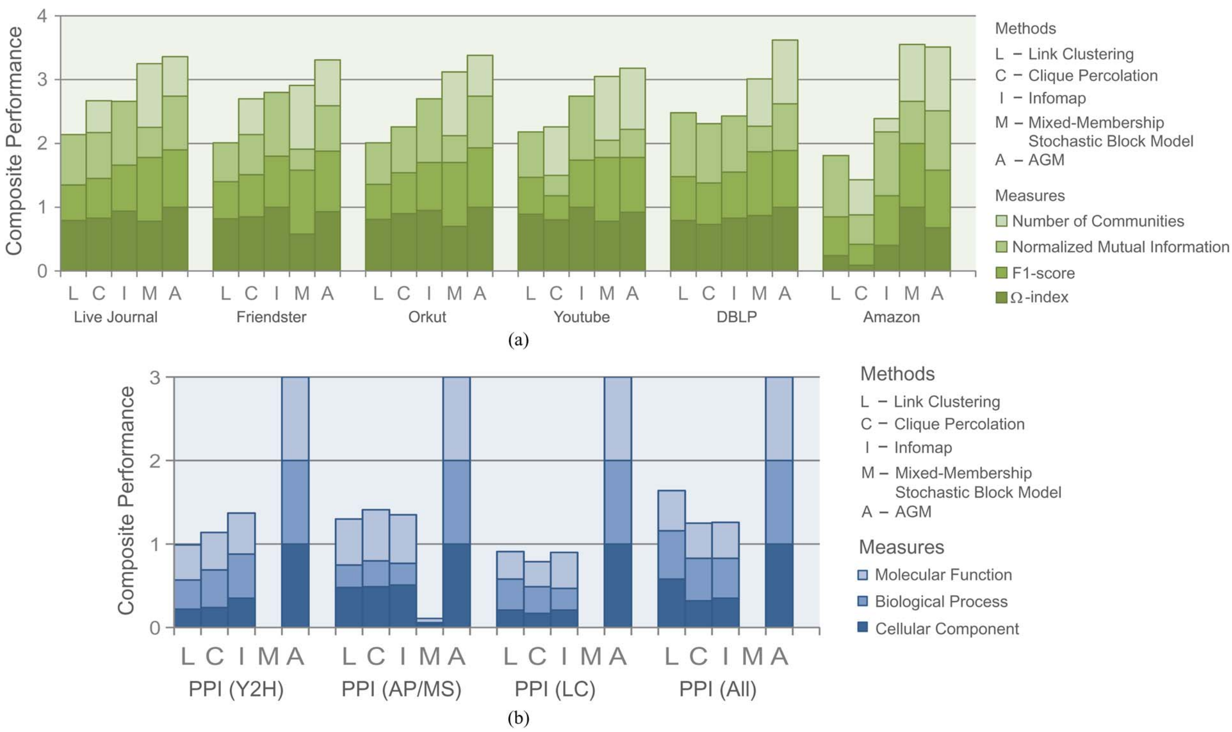


Fig. 11. ***Composite performance of the community detection methods on: (a) six networks with externally labeled ground-truth communities; and (b) four biological networks.***

higher than link clustering (0.38), 22% higher than CPM (0.37), 5% higher than Infomap (0.44), and 26% higher than MMSB (0.36).

In terms of absolute values of scores, AGM archives the average F1 score of 0.57, average omega index of 0.46, mutual information of 0.15, and accuracy of the number of communities 0.42. We also note that AGM also outperforms CPM with other values of k ($k = 3, 4, 6$).

APPENDIX IV
APPLYING AGM TO BIOLOGICAL NETWORKS

We also evaluate the performance of AGM on the four types of PPI networks of *Saccharomyces cerevisiae* [6]. As performance metrics, we compute the average statistical significance of detected communities (p-value) for the three types of GO terms (biological process, cellular component, and molecular function) [41].We consider negative logarithm of average p-values for each of the three GO term types as three separate scores.

Fig. 11(b) displays the composite performance in the four PPI networks. We observe that the AGM attains the best composite performance in all four networks. On average, the composite performance of AGM is 3.00, which is 150% higher than that of link clustering (1.20), 163% higher than that of CPM (1.14), 148% higher than that of Infomap (1.21), and 12 times higher than that of MMSB (0.08). We further investigated the poor performance of MMSB on these networks and found it is due to the fact that MMSB tends to find very large communities consisting of more than 80% of the nodes.

Acknowledgment

The authors would like to thank R. Sosič, P. Mason, M. Macy, S. Fortunato, D. McFarland, and H. Garcia-Molina for invaluable discussions and feedback.

REFERENCES

[1] S. Fortunato, "Community detection in graphs," *Phys. Rep.*, vol. 486, no. 3–5, pp. 75–174, 2010.

[2] N. Krogan *et al.*, "Global landscape of protein complexes in the yeast saccharomyces cerevisiae," *Nature*, vol. 440, no. 7084, pp. 637–643, 2006.

[3] S. Wasserman and K. Faust, *Social Network Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1994.

[4] G. Flake, S. Lawrence, C. Giles, and F. Coetzee, "Self-organization and identification of web communities," *Computer*, vol. 35, no. 3, pp. 66–71, 2002.

[5] M. Newman, *Networks: An Introduction*. Oxford, U.K.: Oxford Univ. Press, 2010.

[6] Y.-Y. Ahn, J. P. Bagrow, and S. Lehmann, "Link communities reveal multi-scale complexity in networks," *Nature*, vol. 466, pp. 761–764, Oct. 2010.

[7] E. M. Airoldi, D. M. Blei, S. E. Fienberg, and E. P. Xing, "Mixed membership stochastic blockmodels," *J. Mach. Learn. Res.*, vol. 9, pp. 1981–2014, 2007.

[8] M. Sales-Pardo, R. Guimerà, A. Moreira, and L. A. N. Amaral, "Extracting the hierarchical organization of complex systems," *Proc. Nat. Acad. Sci. USA*, vol. 104, pp. 18 874–18 874, 2007.

[9] I. Psorakis, S. Roberts, M. Ebden, and B. Sheldon, "Overlapping community

detection using Bayesian non-negative matrix factorization," *Phys. Rev. E*, vol. 83, 2011, 066114.

[10] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, "Uncovering the overlapping community structure of complex networks in nature and society," *Nature*, vol. 435, no. 7043, pp. 814–818, 2005.

[11] T. S. Evans and R. Lambiotte, "Line graphs, link partitions, and overlapping communities," *Phys. Rev. E*, vol. 80, 2009, Art. ID. 016105.

[12] A. Lancichinetti and S. Fortunato, "Community detection algorithms: A comparative analysis," *Phys. Rev. E*, vol. 80, no. 5, 2009, 056117.

[13] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *J. Stat. Mech., Theory Exp.*, vol. 2008, no. 10, 2008, P10008.

[14] M. Newman and M. Girvan, "Finding and evaluating community structure in networks," *Phys. Rev. E*, vol. 69, 2004, 026113.

[15] P. J. Mucha, T. Richardson, K. Macon, M. A. Porter, and J.-P. Onnela, "Community structure in time-dependent, multiscale, and multiplex networks," *Science*, vol. 328, no. 5980, pp. 876–878, 2010.

[16] C. Granell, S. Gómez, and A. Arenas, "Hierarchical multiresolution method to overcome the resolution limit in complex networks," *Int. J. Bifurcat. Chaos*, vol. 22, no. 7, 2012, 1250171.

[17] B. Ball, B. Karrer, and M. E. J. Newman, "Efficient and principled method for detecting communities in networks," *Phys. Rev. E*, vol. 84, 2011, 036103.

[18] J. Yang and J. Leskovec, "Defining and evaluating network communities based on ground-truth communities," in *Proc. IEEE Int. Conf. Data Mining*, 2012, pp. 745–754.

[19] J. Yang and J. Leskovec, "Community-affiliation graph model for overlapping network community detection," in *Proc. IEEE Int. Conf. Data Mining*, 2012, pp. 1170–1175.

[20] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," *Proc. Nat. Acad. Sci. USA*, vol. 105, pp. 1118–1123, 2008.

[21] S. P. Borgatti and M. G. Everett, "Models of core/periphery structures," *Social Netw.*, vol. 21, pp. 375–395, 1999.

[22] P. Holme, "Core-periphery organization of complex networks," *Phys. Rev. E*, vol. 72, 2005, 046111.

[23] F. D. Rossa, F. Dercole, and C. Piccardi, "Profiling core-periphery network structure by random walkers," *Sci. Rep.*, vol. 3, 2013, DOI: 10.1038/srep01467.

[24] J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney, "Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters," *Internet Math.*, vol. 6, no. 1, pp. 29–123, 2009.

[25] A. Clauset, M. Newman, and C. Moore, "Finding community structure in very large networks," *Phys. Rev. E*, vol. 70, 2004, 066111.

[26] M. Coscia, G. Rossetti, F. Giannotti, and D. Pedreschi, "Demon: A local-first discovery method for overlapping communities," in *Proc. ACM SIGKDD Int. Conf. Knowl. Disc. Data Mining*, 2012, pp. 615–623.

[27] J. Leskovec, K. Lang, and M. Mahoney, "Empirical comparison of algorithms for network community detection," in *Proc. Int. Conf. World Wide Web*, 2010, pp. 631–640.

[28] D. Watts and S. Strogatz, "Collective dynamics of small-world networks," *Nature*, vol. 393, pp. 440–442, 1998.

[29] J. A. Davis, "Clustering and Structural Balance in Graphs," *Human Relations*, vol. 20, pp. 181–187, 1967.

[30] M. S. Granovetter, "The strength of weak ties," *Amer. J. Sociol.*, vol. 78, pp. 1360–1380, 1973.

[31] A. Clauset, C. Moore, and M. Newman, "Hierarchical structure and the prediction of missing links in networks," *Nature*, vol. 453, no. 7191, pp. 98–101, 2008.

[32] F. Radicchi, C. Castellano, F. Cecconi, V. Loreto, and D. Parisi, "Defining and identifying communities in networks," *Proc. Nat. Acad. Sci. USA*, vol. 101, no. 9, pp. 2658–2663, 2004.

[33] J. Yang, J. McAuley, and J. Leskovec, "Community detection in networks with node attributes," in *Proc. IEEE Int. Conf. Data Mining*, 2013, pp. 1151–1156.

[34] J. J. Yang, McAuley, and J. Leskovec, "Detecting cohesive and 2-mode communities in directed and undirected networks," in *Proc. ACM Int. Conf. Web Search Data Mining*, 2014, pp. 323–332.

[35] J. Yang and J. Leskovec, "Structure and overlaps of communities in networks," Stanford InfoLab, Tech. Rep., Oct. 2014. [Online]. Available: http://snap.stanford.edu/agm

[36] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annu. Rev. Sociol.*, vol. 27, pp. 415–444, 2001.

[37] G. Simmel, *Conflict: The Web of Group Affiliations, Trans. by Kurt H. Wolff and Reinhard Bendix*. Florence, MA, USA: Free Press, 1955.

[38] R. L. Breiger, "The duality of persons and groups," *Social Forces*, vol. 53, no. 2, pp. 181–190, 1974.

[39] S. Lattanzi and D. Sivakumar, "Affiliation networks," in *Proc. 41st Annu. ACM Symp. Theory Comput.*, 2009, pp. 427–434.

[40] J. Yang and J. Leskovec, "Overlapping community detection at scale: A non-negative factorization approach," in *Proc. ACM Int. Conf. Web Search Data Mining*, 2013, pp. 587–596.

[41] E. Boyle *et al.*, "GO::TermFinder—Open source software for accessing gene ontology information and finding significantly enriched gene ontology terms associated with a list of genes," *Bioinformatics*, vol. 20, no. 18, pp. 3710–3715, 2004.

[42] M. P. Rombach, M. A. Porter, J. H. Fowler, and P. J. Mucha, "Core-periphery structure in networks," *SIAM J. Appl. Math.*, vol. 74, no. 1, pp. 167–190, 2014.

[43] R. E. Ulanowicz, C. Bondavalli, and M. S. Egnotovich, "Network analysis of trophic dynamics in south Florida ecosystem, FY 97: The Florida Bay ecosystem," Annu. Rep. U.S. Geol. Service Biol. Resources Div., 1998, pp. 98–123.

[44] S. Lehmann, M. Schwartz, and L. K. Hansen, "Biclique communities," *Phys. Rev. E*, vol. 78, 2008, 016108.

[45] P. W. Holland, K. B. Laskey, and S. Leinhardt, "Stochastic blockmodels: First steps," *Social Netw.*, vol. 5, no. 2, pp. 109–137, 1983.

[46] C. Lee, F. Reid, A. McDaid, and N. Hurley, "Detecting highly overlapping community structure by greedy clique expansion," in *Proc. 4th SNA-KDD Workshop*, pp. 33–42, 2010.

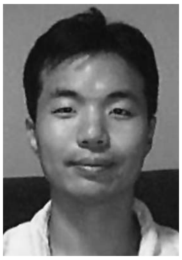
[47] H. Shen, X. Cheng, K. Cai, and M.-B. Hu, "Detect overlapping and hierarchical community structure in networks," *Physica A, Stat. Mech. Appl.*, vol. 388, no. 8, pp. 1706–1712, 2009.

[48] B. Karrer and M. Newman, "Stochastic blockmodels and community structure in networks," *Phys. Rev. E*, vol. 83, 2010, 016107.

[49] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[50] S. Gregory, "Fuzzy overlapping communities in networks," *J. Stat. Mech., Theory Exp.*, vol. 2011, no. 2, 2011, P02017.

ABOUT THE AUTHORS



Jaewon Yang received the B.S. degree in electrical engineering from Seoul National University, Seoul, Korea and the M.S. degree in statistics and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA.

He is a Research Scientist at LinkedIn. His research interests include network community detection, matrix factorization, and recommender systems.

Dr. Yang received the SIGKDD doctoral dissertation award in 2014 and the IEEE ICDM best paper award in 2010.

Jure Leskovec received the B.S. degree in computer science from the University of Ljubljana, Ljubljana, Slovenia, the Ph.D. degree in machine learning from Carnegie Mellon University, Pittsburgh, PA, USA, and postdoctoral training from Cornell University, Ithaca, NY, USA.

He is an Assistant Professor of Computer Science at Stanford University. His research focuses on mining large social and information networks. Problems he investigates are motivated by large scale data, the Web, and online media.

Dr. Leskovec received several awards including a Microsoft Research Faculty Fellowship, the Alfred P. Sloan Fellowship, and numerous best paper awards.

INVITED PAPER

Social Persuasion in Online and Physical Networks

This paper proposes that social persuasion will seamlessly span and use computationally and empirically rigorous methods to understand both the cyber and physical worlds.

By Vivek K. Singh, Ankur Mani, and Alex Pentland

ABSTRACT | Social persuasion to influence the actions, beliefs, and behaviors of individuals, embedded in a social network, has been widely studied. It has been applied to marketing, healthcare, sustainability, political campaigns, and public policy. Traditionally, there has been a separation between physical (offline) and cyber (online) worlds. While persuasion methods in the physical world focused on strong interpersonal trust and design principles, persuasion methods in the online world were rich on data-driven analysis and algorithms. Recent trends including Internet of Things, "big data," and smartphone adoption point to the blurring divide between the cyber world and the physical world in the following ways. Fine grained data about each individual's location, situation, social ties, and actions are collected and merged from different sources. The messages for persuasion can be transmitted through both worlds at suitable times and places. The impact of persuasion on each individual is measurable. Hence, we posit that the social persuasion will soon be able to span seamlessly across these worlds and will be able to employ computationally and empirically rigorous methods to understand and intervene in both cyber and physical worlds. Several early examples indicate that this will impact the fundamental facets of persuasion including who, how, where, and when, and pave way for multiple opportunities as well as research challenges.

KEYWORDS | Cyber–physical social networks; networked intervention; persuasive computing; social persuasion

Manuscript received March 21, 2014; revised September 2, 2014; accepted September 29, 2014. Date of publication November 6, 2014; date of current version November 18, 2014. *(Vivek K. Singh and Ankur Mani contributed equally to this work.)*

V. K. Singh is with the School of Communication and Information, Rutgers University, New Brunswick, NJ 08904 USA and also with the Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: v.singh@rutgers.edu).

A. Mani is with the Stern School of Business, New York University, New York, NY 10012 USA and also with the Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: amani@mit.edu).

A. Pentland is with the Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: pentland@mit.edu).

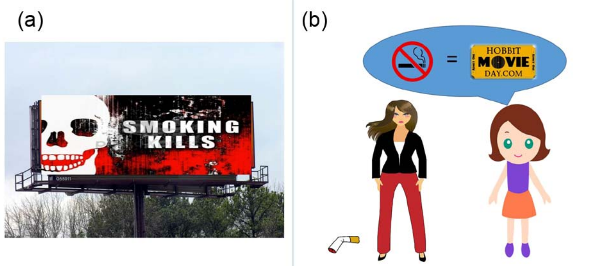
Digital Object Identifier: 10.1109/JPROC.2014.2363986

I. INTRODUCTION

Imagine Alice, a 20-year-old senior in college, trying to quit smoking. She has not smoked in a month. On a Saturday afternoon, she goes alone to the terrace of her dorm with a cigarette and a lighter. Just as she is about to light her cigarette, her friend, Jane, from the adjacent room comes and says, "Stop! I will come with you to watch *The Hobbit* if you do not light that cigarette." Alice does not light the cigarette, and the two friends enjoy a wonderful movie together.

This was not a coincidence. Multiple events took place in the background that allowed Jane to persuade Alice to stop smoking. Alice had signed up for a program to quit smoking. The program collects information about Alice and her friends. Several pieces of information such as location, intent, friendship patterns, and recent actions were monitored. The program recognized that Alice was lonely, because her boyfriend was out of town, and she could not find someone to go watch *The Hobbit* with her. She had reported on her online social network that she is looking for company to go watch the movie. Alice did not get along with her roommate, so when her roommate came to the room, she found an excuse to go to the terrace and smoke. The risk for Alice slipping was very high, so the program recognized that it was the right moment to persuade her to not smoke. Given Alice's location and the availability of her friend, Jane, next door, Jane was the perfect candidate to persuade her. Alice's risk of smoking at the terrace and her intent to watch the movie was communicated to Jane by a mobile app message and that suggested Jane the ideal way to persuade Alice.

Stories of social persuasion like this are going to be very common in future. The persuasion here was optimized for the aspects of *who*, *how*, *when*, and *where*. With the emergence of fine-grained data about users and their social context in the physical (offline) and cyber (online) worlds, always-on sensing, and widespread accessibility of



WHO: Anthropomorphic agent
HOW: Message to scare
WHEN, WHERE: Fixed (a place with high traffic)

WHO: Nearby friend with good ties.
HOW: Social (non-monetary) incentives.
WHEN, WHERE: Situation-aware, just in time

Fig. 1. ***Comparison between (a) traditional and (b) emerging persuasion strategies. Emerging strategies will frequently leverage a user's social ties and positive nonmonetary incentives, and be situation aware.***

enabling technologies, we are stepping into an era of *ex-post* optimization of social persuasion. As shown in Fig. 1, not too long ago, the mechanisms of persuasion for quitting smoking involved banners on highway that said, "Smoking kills." The message and the location were *ex-ante* optimized to persuade the largest population of smokers to not smoke. Today, due to the availability of rich personal, social, and contextual data, similar persuasion attempts can leverage social ties, employ nonmonetary incentives, and be responsive to user situations.

This is possible in large part due to the technological trends including the Internet of Things, mobile phone usage, and mediated human interaction. These trends are paving way for an era where computational systems will break the conventional silos of the physical and cyber web. People's real world movements, habits, and social connections will be accessible via the ubiquitous web, and multiple layers of "cyber" data including information hidden in webpages, databases, and online social networks will be available to apps running on each user's mobile phone. Such apps will be able to integrate heterogeneous data to understand both the spatio–temporal and social contexts, and be able to respond to human needs at the right time, right place, and in the right social context.

These trends will impact the persuasion frameworks being employed. Traditionally, the persuasion framework involving user actions, generated data, and interventions (see Fig. 2) have been siloed, i.e., focused within one realm. For example, in the cyber realm, a user's online search history was used to recommend products and the click through (if undertaken) was tracked. Soon, the computational mechanisms will be able to select the right approach for persuasion which could also be based on combination of the cyber and physical webs. For example, a user's online patterns indicating emotional needs could be intervened by real-world actions by friends and family. Taken together, these methods will allow humans to persuade each other and impact multiple facets of human lives including health, traffic, water, disaster mitigation, epidemic control, financial mechanisms, security, and politics.

In this position paper, we illustrate the emerging technological changes and discuss how they will impact social persuasion in the emerging cyber–physical social networks. We expect the technology to impact the persuasion landscape in multiple important ways: 1) merging of the silos of data; 2) persuasion mechanisms that work in an always-on and just-in-time manner; 3) scale and resolution of the data available to persuasion systems; and 4) the emergence of closed-loop persuasion systems. While such technologies and corresponding methods will impact societies in multiple ways, we scope the discussion here on persuading users individually (rather than *en mass*) via intervention mechanisms that optimize for the essential aspects of who (social), when, where (situational), and how (channel).

The focus of this paper is different from automated intervention mechanisms (alerts, automated reminders) that do not have any other human in the loop. This is because, first, human actors are known to be much more persuasive than anthromorphized agents. Even more importantly, humans can act as a "sounding board" for the advice generated by automated means. Multiple aspects of intervention (ethical, social, and also verificational) are best judged by a fellow human than an entirely automated process. For example, in the smoking scenario, Jane could act as a social filter who could first do a "sanity check" to ensure that Alice might indeed be at risk, and second, consider that watching a movie together is an appropriate and ethically sound method for intervention. Hence, while automated systems will increasingly provide better recommendations, a human in the loop would still be crucial to their impact in real-world social settings. Similarly, a human-in-the-loop intervention is different from changes made by system designers, who are aware of the global network structure, and can manipulate the network structure or the information content without the users realizing it. Such scenarios raise ethical questions, as was seen in the recent response to [18]. This paper instead focuses on scenarios in which there is an explicit action by a human in the loop to persuade his or her peers.

II. CURRENT APPROACHES

There are multiple tools and approaches that are already being applied at large scales in both cyber and physical social networks.

In physical social networks, the importance of social proof and trust has been well documented. For example, Golembiewski and McConkie [12] have argued the case for the importance of trust in mediating social processes. Similarly, Brown and Reingen [5] have reported quantitative

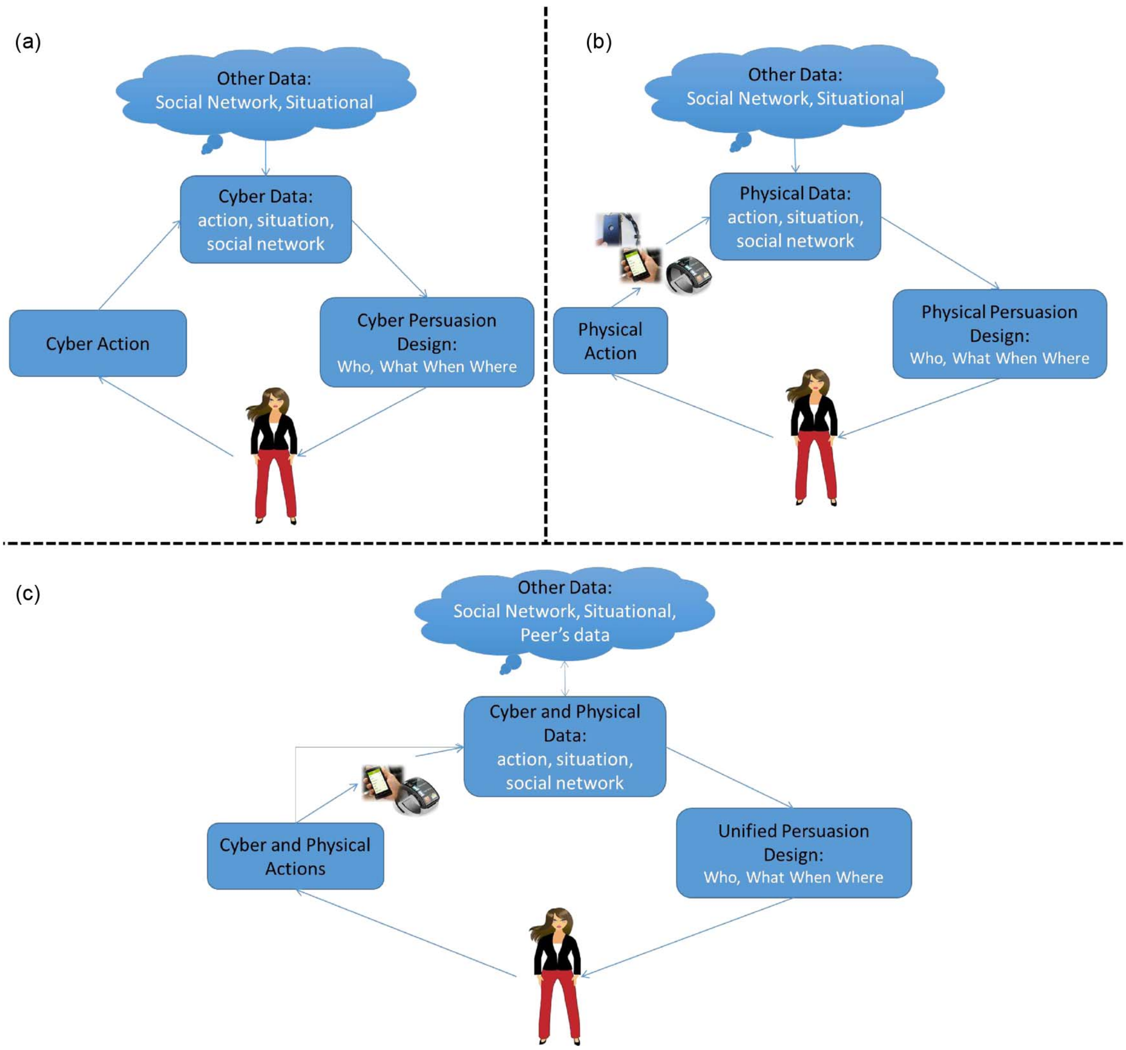


Fig. 2. ***Conventionally, the persuasion scenarios involving user actions, generated data, and persuasion design have focused only on one realm: (a) cyber, or (b) physical. Emerging persuasion scenarios (c) will be able to combine and move seamlessly across the cyber and physical worlds for understanding the actions, capturing the data, and intervening.***

results highlighting the value of word of mouth for marketing campaigns. In the last decade, multiple government and nongovernment agencies have been providing comparison metrics to other users to convince users to regulate their consumption of electricity, water, food, and other sustainability-driven choices. For example, in a study involving more that 80 000 households, Alcott [3] found that social comparison was an effective method for reducing the energy consumption of households.

In cyber social networks (e.g., Twitter and Facebook), users' comments and views on products, brands, and issues are already being viewed as electronic words of mouth [15]. These word-of-mouth impressions can persuade others to adopt or reject certain products or services [13]. Multiple advertising campaigns highlight the users in one's online social network (e.g., Facebook and Google+) who have also bought the same product, consumed the same media content, or taken similar ideological positions. Taken together, the data relating to every "like," "poke," tweets, search history, articles read, media consumed, and messages shared are being analyzed by online firms using data-driven techniques (network analysis, user profiling, influence analysis, contagion and homophily effects, associative rule mining) to recommend products and services to users based on the behavior of other users. For example, the collaborative filtering [21] approach, which recommends products to users based on the ratings provided by "similar" users has been widely adopted by online firms to recommend products and services to users.

Traditionally, while the offline networks have involved a much stronger sense of trust, human intelligence, and real-world context, the online settings have had a richer access to data, algorithms, and analytics. The emergence of "big data," Internet of Things, and similar advancements is changing this scenario. Computing technologies are now able to understand detailed human behavior in physical-world settings. With the right permissions, it is now possible to capture every gaze, interest, and heartbeat of any given user. With mobile phones becoming the key enablers, it is possible to use computational mechanisms and data-driven approaches originally defined for "cyber" networks to now work with physical behavioral data.

III. TECHNOLOGY CHANGES

The Internet of Things has been a huge driver for the merging of the cyber and physical webs. According to Acatec, the German Academy of Technical Sciences, over 98% of microprocessors today are embedded in everyday objects and devices [1]. Similarly, it is estimated that today there are more than 12 billion Internet-enabled devices [8], and more active phone connections than the population of the world. According to Walker Sands (Chicago, IL, USA), over 28% of the Internet traffic requests come from mobile devices [26]. A unique property of these mobile phones is that any data coming from them are inherently spatio–temporal [exact Global Positioning System (GPS) or coarse cell tower and timestamps]. All of these are part of a growing trend. The International Data Corporation (IDC) reported that the number of smartphones sold was already more than the number of "nonsmart" phones in 2013 [14]. This means that devices, which can capture human movements (GPS), face-to-face interactions (via Bluetooth, infra-red, GPS), and call/SMS social networks as they occur in the physical world are soon going to be ubiquitous.

These trends impact the technology in multiple important ways.

1) *Merging silos of data:* While the traditional methods for data-driven persuasion focused on only one type of data (either cyber or physical) in a single format, multiple emerging technologies are enabling combination of these data for a more holistic understanding of the user situation. These technologies include the semantic web, information fusion, federated databases, and mashups, and are being applied for applications ranging from healthcare, to travel, and politics. For example, Pongpaichet *et al.* [25] describe a method for integration of user's personal context with distributed spatio–temporal data to create the right interventions for allergy patients.
2) *Always on, just in time:* While traditional persuasion strategies were employed in limited spatial and temporal bounds, today's computational systems are always on. Apps running on the mobile phone are with the person 24/7, wherever he/she goes. This allows the intervention mechanisms to respond just in time to undertake preventive measures and allow for *ex-post* optimized persuasion.
3) *Scale and resolution:* The emerging trends on "big data" imply that computational systems have access to information at scales and resolution levels that were never captured before. For example, today every gaze, glance, heartbeat, emotion, movement, financial activity, and social activity of a person can be digitally captured and shared with the community if the person chooses so. This implies that systems can be personalized in ways not possible before. Similarly, satellite imagery, Internet-of-Things-based devices, sensor networks, and projects such as the Planetary Pulse are channeling data coming from more parts of planet Earth in more detail than ever before to users and their mobile applications. This, in effect, allows user applications to have access to the pulse of the planet and the actions of the society [11] while taking every action.
4) *Closed-loop systems:* Siloed persuasion strategies were often open loop. For example, it was very hard for online smoking cessation campaigns to follow through and observe the physical actions of the users. Even within the physical realm it was impractical for persuaders (e.g., smoking awareness volunteers) to observe the actions of their subjects. The newer technologies are allowing for the impact of persuasion strategies to be observed in a closed loop. Over time these systems will identify which strategies work best in different scenarios.

These technology changes are also allowing scientists to study social persuasion at newer scales and granularity, and cause *in situ* interventions by combining multiple layers of data. Multiple early initiatives have already started building tools, algorithms, and techniques that employ smart devices to understand and influence cyber–physical social networks.

For example, the "Friends and Family" study conducted at the Media Lab, Massachusetts Institute of Technology (MIT, Cambridge, MA, USA), studied a community of 100+ users living in a residential dorm for a period of over a year [2]. They obtained face-to-face interaction data, Facebook interaction data, as well as self-reported social ties via surveys. In multiple studies, they have shown how face-to-face and other types of networks can be combined to predict flu spread, spending patterns, mobile app adoption, and to encourage users to undertake certain actions such as jogging [2], [22], [28]. A related effort is combining layers of data ranging from Twitter streams and air quality levels to personal GPS coordinates and accelerometer readings to cause just-in-time interventions [25].

Similar effort is being conducted by the University of Trento (Trento, Italy) under the umbrella of "Mobile Territorial Lab" where multiple studies are being conducted to understand user behavior in "work" as well as "personal" social environments. For example, the "Socio-Metric Badges" study analyzed social interaction data for six weeks in a research institution monitoring the interaction activity of 53 people [20]. The generated corpus allows researchers and practitioners with a digital trace data of people's physical as well as online (e-mail) social interaction behavior. With supporting information about participants' individual characteristics (e.g., personality traits) and the interaction context (e.g., participants' current situation), this study is being expanded onto a broader program where a community of 100+ users is being studied in terms of their spending habits and affect levels.

The Copenhagen Networks Study at the Technical University of Denmark (Lyngby, Denmark) [30] has been using smartphones and the associated sensors (GPS, WiFi access points, calls) as well as Facebook messages to understand a community of freshmen at the university. The NetSense project at the University of Notre Dame (Notre Dame, IN, USA) also analyzes the social interaction patterns in a community of 200 freshmen as measured through text, voice call, e-mail, Facebook posts, and the proximity between the devices. Such initiatives point to a growing interest in studying physical social networks at scale: 100+ users and multiple months, and complement it with online networks and contextual data. The "Phone lab" initiative (http://www.phone-lab.org/) at the State University of New York at Buffalo (SUNY Buffalo, Buffalo, NY, USA) provides a public Android testbed designed to simplify large-scale social experiments that can be undertaken via smartphones. Initiatives like these may soon make experimentation and analyses in physical social networks accessible to a much larger pool of researchers and practitioners.

IV. IMPACT ON PERSUASION

These technology changes are blurring the boundaries between online and offline (or cyber and physical) social networks. We expect many of the computationally rigorous methods that were originally designed for the cyber data to evolve to consider the rich contextual data provided by the physical sensors. Specifically looking back at the four key aspects identified in Section I, we expect the systems to be able to understand the who, how, when, and where aspects in much greater detail than possible before.

A. Who

Multiple studies have suggested that identifying the right node for conveying a message is extremely important for successful persuasion [6], [32]. People respond to persuasion by close friends and family as opposed to strangers, and persuasion by people with authority [6], [24]. Our earlier work has also shown that close friends could be very persuasive [2], [22]. Segmentation-based approaches are used to spread messages to a group of similar people, for example, those who share a common passion for rock music or certain sports, or political ideologies. Induction tries to activate newer connections between users where certain thought leaders, celebrities, or early adopters are encouraged to communicate the message and persuade people. This effect is seen also in social media: celebrities are often paid to tweet about products and multiple firms try to make their campaigns go "viral." Last, alteration of networks to change the underlying interconnections is an emerging but extremely powerful mechanism for behavior change. For example, Aharony *et al.* [2] experimented with a social mechanism, described in [23], where the peers of the target users were rewarded rather than the target users themselves. This strategy was found to be more effective at persuading users to exercise than the traditional approach of paying the users themselves. Our previous work has also shown that emerging technologies (smartphones with physical proximity sensors) and computational approaches can also be used to automatically recognize close and trusted ties. In fact, these trusted ties were found to be even more effective at causing behavior change than the close ties [27].

Peer influence for persuasion is more pronounced for products and services with network externalities like phone communication plans and adoption of online social networks such as Facebook. However, the earlier choice of choosing peers and celebrities was *ex-ante* optimized for an assumed distribution about the population without detailed information about peer relationships and individual likings for celebrities.

We have also presented the theoretical underpinnings of this phenomenon. Our results on the joint model of externalities and peer pressure show that even after considering the (positive and negative) changes in the relationship between the two agents in a persuasion scenario, using right peers to persuade can help control global externalities much more efficiently than direct persuasion through subsidies [23]. For example, in the described smoking scenario, there is a cost associated with Jane's persuasion, and it may impact the relationship between the two both positively or negatively. Our model in [23] indicates that using right peers to persuade is more efficient than direct persuasion through subsidies.

Going forward, the systems with an ability to merge data across silos at high scale and resolution in real time will be able to identify the right person to initiate the intervention. Further, the information about these interventions and the success/failure of them in terms of actual user actions could be tracked to refine the social ties as well as strategy scores. Over time, these may allow systems to adapt and also point out relevant trends on the success and failure of various persuasion strategies.

B. How

An important change that technology brings is that there is increased and accurate information about intent and preferences of the individuals needed to be persuaded. In the example, Alice was actually lonely and was going to smoke. The standard pricing mechanisms that would pay Alice a little money to not smoke would not have had a big impact. However, a company to watch her favorite movie was a big incentive for her, worth a lot more than a little amount of money. Persuasion theories have utilized several ways to persuade, such as using force, appealing to reason, appealing to emotion, coercion, and deception [6]. The use of force is considered the failure of persuasion [33]. Public policies such as taxation and subsidies are often designed to appeal to reason, while advertising is often appealing to emotion, coercion, and sometimes deception.

Our earlier work has argued a case for the leveraging the difference between "incurred cost" and "perceived value," especially in nonmonetary transactions [29]. For example, better game armor, "mayor" status, and higher download bandwidth typically cost much less to the enabling platform than their perceived value by the user. Similarly, social incentives can be a lot more effective than purely monetary incentives. In our previous work [23], we found that peer persuasion via payments to friends was 3.5 times more effective at causing behavior change than direct payment to users. In fact, we have also found that passive social persuasion can already be effective in multiple application settings. For example, a previous study in the group on Meeting Mediator—a mobile system that detects social interactions and provides real-time feedback to enhance group collaboration and performance—showed that visualizing the social interaction pattern data in real time on the mobile phone of each user could induce changes in group collaboration patterns [17]. In particular, the results show greater productivity and trust within geographically separated groups that are using the Meeting Mediator. A different study conducted by Balaam *et al.* [4] used a multiuser public display to enhance the interactional synchrony by visualizing subtle feedback about users' behavior. Their results suggest that social dynamics can be used by machines to support group behavior without requiring a direct and exclusive interaction with the users.

Since one technique often does not fit all people, the emerging trends of fine-grained information about individual preferences can help not just identify the optimal method but also what will appeal to the individual the most and how to persuade can be *ex-post* optimized as well. The merging of online and offline worlds also creates possibilities to provide incentives to people in the physical world for actions in online worlds. Often people are given discount coupons to restaurants for taking an online survey. Several such possibilities are being increasingly made possible by the virtual currencies such as bitcoin.

C. When and Where

An understanding of the user situation allows the system to intervene at the most opportune time and place. For example, the intervention by Jane in the example in Section I at the "right" time and place was critical to its success. The relationship between time and place gives a good estimate of point of action, and persuasion is very effective at the point of taking action. The timing of the intervention has been identified to be a critical determinant of success in Fogg's behavioral persuasion model [9], and similar results have been reported in practical intervention studies in interpersonal settings [34]. The timing and the location are important aspects for the success of geofencing-based approaches for marketing and advertisements. Users are more likely to be interested in discount coupons or physically be able to attend shows and concerts when they are in the vicinity to these establishments. Pushing upgrades, up-sells, and checkout-counter purchases have been well documented in terms of their effect on purchase behavior. These approaches also connect very well with the "bait-and-switch" or the "commitment-and-consistency" principle proposed by Cialdini [6].

Some of our recent work has focused on providing users the right situational interventions just when and where they need them. For example, Pongpaichet *et al.* [25] define a generic approach for users to receive allergy/asthma related alerts just as the combination of their personal and spatio–temporal parameters matches certain criteria. The approach of intervening at the right time and place has also been adopted by multiple other efforts. For example, multiple studies have shown that the placement and display of water meters right when one is taking the shower can be a lot more effective than posteffect awareness [16], [31].

The emerging always-on technologies that are able to cause the right "situational intervention" at the right time will allow future systems to monitor and maybe even predict the right time to initiate an intervention. In fact, Google "Now" is providing anticipatory methods to send alerts to users about things that maybe of interest to them in the near future. For example, if a person has already booked and paid for a hunting trip, it will be difficult to convince her to not go for the trip as she is leaving her home. However, if the peer of a person was available to persuade the person (online or in person) at the time of purchase of the trip, then the persuasion will be more effective. The technological changes will also help identify such persuasion opportunities and make such persuasions possible.

V. RESEARCH OPPORTUNITIES

The intersection of the online and offline social networks creates multiple novel opportunities to devise tools, techniques, and algorithms that connect varied information and persuasion channels across these networks. While many of the existing research directions will need to be reexamined and refined to support for this intersection, certain newer challenges will become exceedingly relevant.

1) *Privacy and ethics:* While privacy and ethics of persuasion were already important concerns in the online networks, the emergence of technology that captures rich personal behavioral (every heartbeat, gaze, interest, mobility pattern) and social interaction (face-to-face interaction, calls, sms, colocation) data and uses them for *in situ* persuasion opens doors to a very different level of ethical and privacy concerns. While users are presumably able to adopt newer cyber identities, physical identities and health parameters once compromised cannot be restored. Hence, the recording and analysis of physical data at the same level of discourse as online data poses multiple privacy risks and hence research challenges. One possible approach to tackle this might lie in creating trusted "personal data stores" [7] that allow for question-and-answer approaches that support such persuasion frameworks without giving away raw data to third parties. Further, a technological ability to persuade does not imply that persuasion should actually be carried out. For example, while many people might support sharing of such information for well-accepted societal goals (e.g., to eradicate behavioral diseases like diabetes, or trigger early interventions to avoid traffic accidents), a much more nuanced discussion is required on the right policies for recommending newer products and commercial services. Clearly, newer research efforts are needed to define the right norms and policies that govern the use of persuasion in cyber–physical social settings. In fact, we anticipate that the same kind of computational mechanisms that have been employed for better "product" recommendations will be adapted to provide "privacy" recommendations to a large number of users.
2) *Orchestration and tradeoffs between cyber and physical persuasion:* So far, the persuasion approaches have stayed within their respective realms (online or physical). Soon the merging of the realms will open up interesting tradeoff and coordination challenges. For example, how many online signatures on an issue at Change.org are as effective as ten people physically protesting about the same issue? Similarly, if both online and physical methods are available for persuasion, which method should be used for which tasks? For example, certain sensitive or health-related campaigns might work best in semianonymized settings, while others will benefit from the trusted ties between users. Further, if certain campaigns require a combination of online and physical intervention, what should be their count and order? While early studies such as [10] have started exploring these issues, many more such efforts are needed.
3) *Living labs for social science:* The emergence of platforms for cyber–physical mining of social behavior and interventions opens the doors to an exciting opportunity to test, validate, and refine multiple social science theories. Multiple social science theories have been based on experiments conducted in limited laboratory settings and self-reported surveys. These approaches were costly, piecemeal, retrospective, and often suffered from perception bias. Hence, an ability to conduct longitudinal studies on social behavior as human beings live their natural lives is emerging as a vital tool for computational social scientists [19]. Further, the opportunity to cause interventions and make changes in these longitudinal studies may allow social scientists to differentiate between correlations and causations and develop normative social science that can potentially improve the quality of human life. ■

REFERENCES

[1] Acatec, *Cyber-Physical Systems: Driving Force for Innovations in Mobility, Health, Energy and Production*. New York, NY, USA: Springer-Verlag, 2012.

[2] N. Aharony, W. Pan, C. Ip, I. Khayal, and A. Pentland, "Social FMRI: Investigating and shaping social mechanisms in the real world," *Perv. Mobile Comput.*, vol. 7, no. 6, pp. 643–659, 2011.

[3] H. Allcott, "Social norms and energy conservation," *J. Public Econ.*, vol. 95, no. 9, pp. 1082–1095, 2011.

[4] M. Balaam, G. Fitzpatrick, J. Good, and E. Harris, "Enhancing interactional synchrony with an ambient display," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2011, pp. 867–876.

[5] J. J. Brown and P. H. Reingen, "Social ties and word-of-mouth referral behavior," *J. Consumer Res.*, vol. 14, no. 3, pp. 350–362, 1987.

[6] R. B. Cialdini, *Influence: The Psychology of Persuasion (Collins Business Essentials)*. New York, NY, USA: Harper Paperbacks, 2007.

[7] Y.-A. de Montjoye *et al.*, "On the trusted use of large-scale personal data," *IEEE Data Eng. Bull.*, vol. 35, no. 4, pp. 5–8, 2012.

[8] D. Evans, "The internet of things: How the next evolution of the internet is changing everything," CISCO White Paper, 2011.

[9] B. J. Fogg, *Persuasive Technology: Using Computers to Change What We Think and Do*. San Mateo, CA, USA: Morgan Kaufmann, 2002.

[10] Georgetown University Center for Social Impact Communication and Waggener Edstrom, "Digital persuasion: How social media motivates action and drives support," Worldwide Report. [Online]. Available: http://waggeneredstrom.com/what-we-do/social-innovation/report-digital-persuasion/

[11] F. Giannotti *et al.*, "A planetary nervous system for social mining and collective awareness," *Eur. Phys. J. Special Top.*, vol. 214, no. 1, pp. 49–75, 2012.

[12] R. T. Golembiewski and M. McConkie, "The centrality of interpersonal trust in group processes," in *Theories of Group Process*, C. L. Cooper, Ed. New York, NY, USA: Wiley, 1975.

[13] T. Hennig-Thurau, C. Wiertz, and F. Feldhaus, "Exploring the 'Twitter effect:' An investigation of the impact of microblogging word of mouth on consumers' early adoption of new products," *Soc. Sci. Res. Netw.*, 2012, DOI: 10.2139/ssrn.2016548.

[14] International Data Corporation (IDC), "Worldwide quarterly mobile phone tracker," 2013. [Online]. Available: http://www.idc.com/tracker/showproductinfo.jsp?prod_id=37

[15] B. J. Jansen, M. Zhang, K. Sobel, and A. Chowdury, "Twitter power: Tweets as electronic word of mouth," *J. Amer. Soc. Inf.*

Sci. Technol., vol. 60, no. 11, pp. 2169–2188, 2009.

[16] K. Kappel and T. Grechenig, "Show-me: Water consumption at a glance to promote water conservation in the shower," in *Proc. 4th Int. Conf. Persuasive Technol.*, 2009, DOI: 10.1145/1541948.1541984article 26.

[17] T. Kim, A. Chang, L. Holland, and A. Pentland, "Meeting mediator: Enhancing group collaboration using sociometric feedback," in *Proc. ACM Conf. Comput. Supported Cooperative Work*, 2008, pp. 457–466.

[18] A. D. Kramer, J. E. Guillory, and J. T. Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 24, pp. 8788–8790, 2014, DOI: 10.1073/pnas.1320040111.

[19] D. Lazer *et al.*, "Life in the network: The coming age of computational social science," *Science*, vol. 323, no. 5915, pp. 721–723, 2009.

[20] B. Lepri *et al.*, "The sociometric badges corpus: A multilevel behavioral dataset for social behavior in complex organizations," in *Proc. IEEE Int. Conf. Privacy Security Risk Trust/Int. Conf. Social Comput.*, 2012, pp. 623–628.

[21] G. Linden, B. Smith, and J. York, "Amazon.com recommendations: Item-to-item collaborative filtering," *IEEE Internet Comput.*, vol. 7, no. 1, pp. 76–80, Jan./Feb. 2003.

[22] A. Madan, K. Farrahi, D. Gatica-Perez, and A. S. Pentland, "Pervasive sensing to model political opinions in face-to-face networks," *Pervasive Computing*, vol. 6696, Berlin, Germany: Springer-Verlag, 2011, pp. 214–231.

[23] A. Mani, I. Rahwan, and A. Pentland, "Inducing peer pressure to promote cooperation," *Sci. Rep.*, vol. 3, 2013, DOI:10.1038/srep01735.

[24] S. Milgram, "Behavioral study of obedience," *J. Abnormal Social Psychol.*, vol. 67, no. 4, pp. 371–378, 1963.

[25] S. Pongpaichet, V. K. Singh, R. Jain, and A. S. Pentland, "Situation fencing: Making geo-fencing personal and dynamic," in *Proc. 1st ACM Int. Workshop Personal Data Meets Distrib. Multimedia*, 2013, pp. 3–10.

[26] Walker Sands Digital, "Mobile traffic report q3 2013," 2013. [Online]. Available: http://www.walkersandsdigital.com/Walker-Sands-Mobile-Traffic-Report-Q3-2013

[27] E. Shmueli, V. K. Singh, B. Lepri, and A. Pentland, "Sensing understanding and shaping social behavior," *IEEE Trans. Comput. Social Syst.*, vol. 1, no. 1, pp. 22–34, Mar. 2014.

[28] V. K. Singh, L. Freeman, B. Lepri, and A. S. Pentland, "Classifying spending behavior using socio-mobile data," *Human*, vol. 2, no. 2, pp. 99–111, 2013.

[29] V. K. Singh, R. Jain, and M. S. Kankanhalli, "Motivating contributors in social media networks," in *Proc. 1st SIGMM Workshop Social Media*, 2009, pp. 11–18.

[30] A. Stopczynski *et al.*, "Measuring large-scale social networks with high resolution," working paper, 2014. [Online]. Available: http://arxiv.org/abs/1401.7233

[31] V. Tiefenbeck, V. Tasic, L. Goette, and T. Staake, "The power of real-time feedback: Evidence from randomized trials with 50 000 showers," presented at the Behav. Energy Climate Change Conf., Sacramento, CA, USA, Nov. 2013.

[32] T. W. Valente, "Network interventions," *Science*, vol. 337, no. 6090, pp. 49–53, 2012.

[33] A. N. Whitehead, "From force to persuasion," in *Adventures of Ideas*, vol. 9317, ch. 5. New York, NY, USA: Simon & Schuster, 1967.

[34] A. W. Woolley, "Effects of intervention content and timing on group task performance," *J. Appl. Behav. Sci.*, vol. 34, no. 1, pp. 30–46, 1998.

ABOUT THE AUTHORS

Vivek K. Singh received the B.S. and M.S. degrees in computer science from the National University of Singapore, Singapore, in 2002 and 2005, respectively, and the Ph.D. degree in information and computer science from the University of California, Irvine, Irvine, CA, USA, in 2012.

He is an Assistant Professor at the Library and Information Science Department, School of Communication and Information, Rutgers University, New Brunswick, NJ, USA and a Visiting Professor at the Media Lab, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. Before joining Rutgers, he was a Postdoctoral Researcher at the MIT Media Lab, and was a full-time Lecturer at the Institute of Technical Education, Singapore, for four years. His research interests lie at the intersection of big data, computational social science, and multimedia information systems.

Prof. Singh was selected as one of the Emerging Leaders in Multimedia Research by IBM Research Labs in 2009, and he recently won the Big Data for Social Good datathon organized by Telefonica, the Open Data Institute, and the MIT.

Ankur Mani received the Ph.D. degree in media, arts and sciences from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2013. His doctoral dissertation was titled "Bilateral exchanges in social networks and the design of public institutions."

He is a Research Scientist at the Information, Operations and Management Sciences Department, Stern School of Business, New York University (NYU), New York, NY, USA and a Visiting Scientist at the Media Lab, MIT. His research lies at the intersection of several disciplines, including computer science, operation research, economics, and sociology. In particular, he focuses on analysis and design of networked sociotechnical systems. His research has been supported by several awards including the Yahoo Key Scientific Challenges, Martin Family Fellowship, and research grants from IBM Research and Natura.

Alex Pentland received the B.S. degree from the University of Michigan, Ann Arbor, MI, USA, in 1976 and the Ph.D. degree from the Massachusetts Institute of Technology, Cambridge, MA, USA, in 1982.

He directs the Human Dynamics Laboratory and the Media Lab Entrepreneurship Program, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA; coleads the World Economic Forum Big Data and Personal Data initiatives; and is a Board member for Nissan, Motorola Mobility, Telefonica, and *Harvard Business Review*. He has previously helped create and direct the MIT's Media Laboratory, the Media Lab Asia laboratories at the Indian Institutes of Technology, and the Strong Hospitals Center for Future Health. His most recent book is *Social Physics* (Baltimore, MD, USA: Penguin Press, 2014).

Dr. Pentland was named by *Forbes* as one of the "seven most powerful data scientists in the world" in 2012, along with Google founders and the CTO of the United States. In 2013, he won the McKinsey Award from *Harvard Business Review*. He is among the most cited computational scientists in the world, and a pioneer in computational social science, organizational engineering, wearable computing (Google Glass), image understanding, and modern biometrics.

INVITED PAPER

Words on the Web: Noninvasive Detection of Emotional Contagion in Online Social Networks

This paper reviews and discusses a nonexperimental, noninvasive method to detect and quantify contagion of semantic expression in massive online social networks.

By Lorenzo Coviello, *Student Member IEEE*, James H. Fowler, and Massimo Franceschetti, *Senior Member IEEE*

ABSTRACT | Does semantic expression spread online from person to person? And if so, what kinds of expression are most likely to spread? To address these questions, we developed a nonexperimental, noninvasive method to detect and quantify contagion of semantic expression in massive online social networks, which we review and discuss here. Using only observational data, the method avoids performing emotional experiments on users of online social networks, a research practice that recently became an object of criticism and concern. Our model combines geographic aggregation and instrumental variables regression to measure the effect of an exogenous variable on an individual's expression and the influence of this change on the expression of others to whom that individual is socially connected. In a previous work, we applied our method to the emotional content of posts generated by a large sample of users over a period of three years. Those results suggest that each post expressing a positive or negative emotion can cause friends to generate one to two additional posts expressing the same emotion, and it also inhibits their use of the opposite emotion. Here, we generalize our method so it can be applied to contexts different than emotional expression and to different forms of content generated by the users of online platforms. The method allows us to determine the usage of words in the same semantic category spread, and to estimate a signed relationship between different semantic categories, showing that an increase in the usage of one category alters the usage of another category in one's social contacts. Finally, it also allows us to estimate the total cumulative effect that a person has on all of her social contacts.

KEYWORDS | Influence; instrumental variables; nonexperimental methods; semantic expression; social networks

I. INTRODUCTION

In the last decade, the challenge of understanding the spreading and synchrony of human behavior over social networks has attracted the attention of the research community at large. The problem originally arises in the context of the social sciences, but due to the expanding usage of online social networks, it has also attracted the interest of the engineering community with the aim of quantifying these effects using the massive amount of data that these networks generate. Studies have included the diffusion of news and "memes" [1]; cascades in communication platforms, networked games, microblogging services [2]; health-related phenomena such as obesity and smoking [3], [4]; emotional states like happiness and depression [5], [6]; purchase of online products [7], [8]; clicking online advertisements; and joining online recreational leagues and store purchases [9].

Manuscript received March 2, 2014; revised October 24, 2014; accepted October 28, 2014. Date of publication November 13, 2014; date of current version November 18, 2014. This work was supported in part by the Army Research Office (ARO) under Grant W911NF-11-1-0363.

L. Coviello and **M. Franceschetti** are with the Department of Electrical and Computer Engineering, University of California San Diego, San Diego, CA 92092 USA (e-mail: lcoviell@ucsd.edu).

J. H. Fowler is with the Department of Political Science and School of Medicine, Medical Genetics Division, San Diego, CA 92093 USA.

Digital Object Identifier: 10.1109/JPROC.2014.2366052

Studies based on observational data pose an inherent difficulty for causal inference because social contacts may have similar behavior as a result of at least two processes: homophily (the tendency of similar individuals to group together) or influence [10], [11]. Controlled experiments allow us to disentangle influence effects from homophily both in the laboratory [12] and online [13]–[16], but they are often limited in scale and lack external validity. Large scale experiments have been shown to be feasible in the context of political participation [16], product adoption [7], [8] and emotional influence [17], but are often impractical or require very close collaboration with private companies.

Moreover, the experimental change in the users' experience required by some of these studies recently came under scrutiny because of questions about the ethics involved. Some people criticized [18] a large scale study [17] of emotional contagion on Facebook in which the researchers changed the content shown to some users in order to study their reaction. Similar criticisms were directed at the online dating website OkCupid for experimenting with their platform in order to understand how individuals react to each other [19]. These recent events call for the development of alternative, nonexperimental methods to study human behavior at large scale [20].

Our work in [21] was an attempt to compensate for the shortcomings of existing experimental and observational approaches, using a method to detect and quantify influence via instrumental variable regression. We studied text-based expression in massive social networks, developed a model of emotional contagion of semantic expression, and validated it on the content posted by a large sample of Facebook users over a period of three years.

In this paper, we show how our model can also be applied to different and possibly heterogeneous data from other social networking platforms, and to contexts other than emotional expression. Our approach is fully nonexperimental: it is based only on observational data and, as a result, it does not alter users' experience. It also guarantees respect for user privacy: for our study in [21], individuals' information and posts were never visible to researchers and resided on secure servers where Facebook stores user data, and were analyzed only at an aggregate level. The study was reviewed for ethics and approved in advance by the Institutional Review Board at the University of California San Diego, San Diego, CA, USA.

Focusing on the mathematical model and on the engineering methodology employed, this paper reviews and complements our previous work. Our individual-level model assumes that a person's usage of words in a semantic category is a linear function of temporal and individual baseline effects; exogenous variables like news, the stock market, or the weather; and endogenous variables—corresponding to the usage of given semantic categories in posts written by the person's social contacts, referred to as "friends." The reciprocal causality between the endogenous variables of the model makes it difficult to obtain consistent and unbiased estimates of social influence. Therefore, we proceed in two steps. First, we aggregate the model on a geographical basis by averaging over all people who are in the same city, obtaining a model based on the same coefficients as the individual-level model but with a much smaller number of observations. Second, we deal with the problem of reciprocal causality by estimating the model using instrumental variable regression, a method pioneered in economics [28]. This method relies on the availability of an exogenous variable—called an instrument—that affects the endogenous variables (friends' posts) but does not directly induce a change in the subject's posts, called the dependent variable. In general, valid instruments might be unavailable, or they might lack sufficient power to predict changes in the endogenous variable. In our work, we considered rainfall experienced by friends as the instrument, using data made available by the National Climatic Data Center (NCDC),[1] which proved to be a robust predictor of emotional expression. Upon finding a relationship between friend's rainfall and their expression, we can assume the former affects the latter as the opposite direction is unlikely. Our method first computes the effect that friends' rainfall (the instrument) has on friends' posts (the endogenous variables). Then, it evaluates the corresponding effect of the rainfall-induced change in friends' expression on the person's posts (the dependent variable).

In order to obtain consistent estimates, the instrument must satisfy the exclusion restriction [28]. This posits that, controlling for all other variables, the instrument (friends' rainfall) must not directly affect the dependent variable. An implication of this restriction is that the instrument must also be uncorrelated with the exogenous variable experienced by the subject (subject's rainfall), otherwise the model might only be estimating how a subject's rainfall affects her own expression. Therefore, to break any correlation between a subject's rainfall and friends' rainfall, we restricted our analysis to observations for which it did not rain in the subject's city. Once this is applied, the subject's rainfall is constant in the data set and, therefore, it does not correlate with friends' rainfall. Moreover, breaking the correlation between user's and friends' rainfall solves the potential issue of the geographic similarity of the weather in close-by cities. As a result, we must also focus exclusively on social ties between individuals in different cities (see Section III-D). Note that individuals in different cities likely do not interact face to face, but they can reach each other via multiple communication media, such as the telephone, e-mail, and social networking websites. Therefore, any influence detected between them is unlikely to be caused by physical

[1]http://www.ncdc.noaa.gov

interaction and would suggest that remote communication plays an important role in spreading semantic expression.

Our method allows us to determine what semantic categories are susceptible to influence between social contacts by estimating how an individual's usage of a semantic category is affected by her friends' usage of the same category. We can then use the estimates for each semantic category to rank them from the most to the least likely to spread.

Moreover, our method allows us to determine the relationship between different semantic categories, by estimating how an individual's usage of one category is altered by her friends' usage of a different category. This will help us to understand whether the usage of a semantic category fosters or inhibits the usage of other categories. We already showed in [21] that expression of positive affect inhibits expression of negative affect and *vice versa*.

Finally, our model allows us to compute the cumulative effect a person has on her friends (see Section III-F). Although the effect on any one social contact will be small, each person typically has many social contacts, so the total expected effect of a single act of expression may alter the expression of several other people. Here, we show how to use our model to quantify this multiplier effect on posts within the same semantic category and on posts in different categories.

A. Related Work

Our work is related to a growing body of literature on influence and diffusion in networks, whose goal is to characterize how behaviors and information spread from person to person. Online social networks are becoming increasingly popular as research environments and sources of data for these investigations. For example, the content posted by people online has been used to identify which people or topics are influential in social networking websites [29] and in the blogosphere [30]. It has also been used to study which network attributes and sharing behaviors make people influential [31], which topics (e.g., represented by hasthtags) diffuse in a more persistent way [32], and even to study the structure of diffusion cascades on different communication platforms [2]. Large scale experimental studies have isolated the role of the network in the diffusion of information [33], emotional expression [17], and behaviors [7], [16]. However, homophily has been shown to play a similarly important role, and scholars have devoted their attention to distinguishing between the two phenomena and to comparing the size of their effects [11], [14], [34], [35].

Our work is related to the econometric literature on instrumental variables. Instrumental variables have been proposed as a tool to infer causal effects from observational data [28]. This approach has been applied to a variety of contexts, such as labor economics [36], the study of the causal effect of education on earning [37], program evaluation [38], the characterization of neighborhood effects [39], and the impact of microfinance [40]. However, valid instruments can be difficult to find [41], and scholars have warned against the risks of using "weak" instruments that do not predict variation in the endogenous variable with sufficient precision [42].

A large body of research studies text meaning by analyzing patterns of words or grammar [43]–[45]. However, the performance of most traditional classification methods relies on sufficient text length, as in the case of bag of words or kernel-based methods [46], [47]. The analysis of short text from microblogging services (such as Twitter or Facebook) requires new approaches [48]–[50], which in some cases leverage metadata (e.g., user's information) or the content of related posts.

Although we mainly focus on the engineering aspects of the detection and measurement of peer influence in semantic expression, our work is also related to sociolinguistics. The full understanding of language in a society requires us to consider the social network in which the language is embedded, intended as the set of relationships and interactions between its individuals [51]. Scholars have argued that speech patterns might depend on the looseness and tightness of the social network [52]. Our model formulation allows us to take tie strength between individuals into account. Different approaches have been proposed to quantify tie strength in online social networks [53], [54], and future research should investigate whether strong ties play a major role in the spread of semantic expression.

II. MODEL VARIABLES

We consider a set T of distinct days. For each day $t \in T$, let $S(t)$ be the population on day t, and let $n(t) = |S(t)|$ be their number. To apply our method, we assume that individuals can be geolocated at the level of cities. For each city g let $S_g(t)$ be the set of individuals in city g on day t and let $n_g(t) = |S_g(t)|$. In general, one might consider different time and geographic resolution. We assume resolution at the level of days and cities in accordance to our previous work [21].

A. Quantifying the Semantic of Text-Based Expression

Several methods can be used to quantify semantic expression of the content posted by individuals (see discussion in Section V). We referred to the semantic categories defined by the Linguistic Inquiry and Word Count (LIWC) 2007 [22], a word classification tool widely used in the social sciences and in psychology research [23]–[27]. The LIWC contains several classes of processes, each of which contains one or more semantic categories, pertaining to affective processes, perceptual processes, biological processes, social processes, and personal concerns. A list of semantic categories from the LIWC is given in Table 1. In [21], we considered the categories for positive and negative affective processes. In general, a

Table 1 List of Semantic Categories From the LIWC

Category name	Example words	Number of words
Social processes	Mate, talk, they, child	455
Family	Daughter, husband, aunt	64
Friends	Buddy, friend, neighbor	37
Humans	Adult, baby, boy	61
Affective processes	Happy, cried, abandon	915
Positive emotion	Love, nice, sweet	406
Negative emotion	Hurt, ugly, nasty	499
Anxiety	Worried, fearful, nervous	91
Anger	Hate, kill, annoyed	184
Sadness	Crying, grief, sad	101
Cognitive processes	cause, know, ought	730
Insight	think, know, consider	195
Causation	because, effect, hence	108
Discrepancy	should, would, could	76
Tentative	maybe, perhaps, guess	155
Certainty	always, never	83
Inhibition	block, constrain, stop	111
Inclusive	And, with, include	18
Exclusive	But, without, exclude	17
Perceptual processes	Observing, heard, feeling	273
See	View, saw, seen	72
Hear	Listen, hearing	51
Feel	Feels, touch	75
Biological processes	Eat, blood, pain	567
Body	Cheek, hands, spit	180
Health	Clinic, flu, pill	236
Sexual	Horny, love, incest	96
Ingestion	Dish, eat, pizza	111
Relativity	Area, bend, exit, stop	638
Motion	Arrive, car, go	168
Space	Down, in, thin	220
Time	End, until, season	239
Personal Concerns		
Work	Job, majors, xerox	327
Achievement	Earn, hero, win	186
Leisure	Cook, chat, movie	229
Home	Apartment, kitchen, family	93
Money	Audit, cash, owe	173
Religion	Altar, church, mosque	159
Death	Bury, coffin, kill	62

larger set C of semantic categories can be considered by our method.

For day $t \in T$ and subject $i \in S(t)$, let $U_i(t)$ be the set of all content posted by subject i on day t, and let $u_i(t) = |U_i(t)|$ be its cardinality. For each subject i such that $u_i(t) > 0$, and each category $c \in C$, let $u_i^{(c)}(t)$ be the number of elements of $U_i(t)$ containing at least one word from category c, and let

$$y_i^{(c)}(t) = \frac{u_i^{(c)}(t)}{u_i(t)}$$

be the frequency, or usage, of category c by subject i on day t. Note that $0 \leq y_i^{(c)}(t) \leq 1$. Therefore, a subject i such that $u_i(t) > 0$ is characterized by $|C|$ variables $y_i^{(c)}(t)$ quantifying her usage of words form all categories in C during day t. Observe that a single piece of user content can contribute to the frequency $y_i^{(c)}(t)$ for several categories c.

B. Exogenous Control Variable

Our method relies on the availability of an exogenous variable that affects the semantic expression of a person's friends but not (directly) the semantic expression of the person. We call this variable the "instrument." Our model characterizes how a change in the instrument induces a change in friends' semantic expression, and how the induced change predicts a change in the person's semantic expression.

There are many sources of exogenous variation in the world, but we chose rainfall as the instrument, relying on data from NCDC. For each city g, we consider the NCDC station closest to it, and let $\bar{x}_g(t) = 1$ if that station recorded rainfall on day t, and zero otherwise. For each subject $i \in S_g(t)$, let $x_i(t) = \bar{x}_g(t)$, that is, a binary indicator variable of rainfall in city g. We focus on rainfall as the instrument for several reasons. First, its geographical resolution lends itself to the analysis of our geographically aggregated model. Second, individuals in the same city tend to experience the same weather on a given day. Moreover, in [21], we show it is a robust instrument in the sense that it captures enough variation of the endogenous explanatory variable (friends' emotional expression). Other meteorological variables would have been a valid alternative. The identification of valid instruments is challenging and finding a systematic way to characterize them is key to apply our method to more general contexts.

C. Social Network Information

For each day $t \in T$, and subjects $i, j \in S(t)$, let $a_{i,j}(t) \in [0, 1]$ be the strength of the relationship from i to j on day t, which need not be symmetric. Also, let $\delta_i(t) = \sum_{j \in S(t)} a_{i,j}(t)$. In [21], we let $a_{i,j}(t) \in \{0, 1\}$, where $a_{i,j}(t) = 1$ denotes that i and j were friends on day t. In this case, $\delta_i(t)$ is the degree of subject i on day t (that is, the total number of friends of the subject). Allowing $a_{i,j}(t)$ to have any value between zero, one would allow to asses the role of tie strength.

III. MODEL

A. Individual-Level Model

Recall that $y_i^{(c)}(t)$ represents the usage of category c by subject i on day t. We assume that $y_i^{(c)}(t)$ is a function of several terms, according to

$$y_i^{(c)}(t) = \theta(t) + f_i + \beta_{c',c} x_i(t) + \gamma_{c',c} \frac{1}{\delta_i(t)} \sum_j a_{i,j}(t) y_j^{(c')}(t) + \epsilon_i(t). \quad (1)$$

$\theta(t)$ represents a "fixed effect" for day t and takes into account temporal patterns of variation in the use of category c (for example, people might be more likely to write about work during the weekdays, or more likely to write about health in the winter). f_i represents a fixed effect for subject i and takes into account different baseline usage of category c for different people (for example, some people might write about work more than others). $x_i(t)$ represents the exogenous variables experienced by subject i on day t. Equation (1) assumes that the effect of the exogenous variable $x_i(t)$ is weighted by a coefficient $\beta_{c',c}$ (the same for all subjects i), whose sign and strength represent the effect of the exogenous variable on usage of semantic category c. The summation in (1) represents the effect of usage of semantic category $c' \in C$ by i's friends on i's usage of category c.[2] Note that the effect of friends' expression is assumed to be inversely proportional to i's degree $\delta_i(t)$, compatible with the idea that a person with a lot of friends is less likely to view posts by all of them. This endogenous term is weighted by the coefficient $\gamma_{c',c}$, which represents the direction and strength of influence (assumed to be the same for all subjects). Finally, $\epsilon_i(t)$ are assumed independent and identically distributed normal error terms with zero mean and variance σ^2, to take unobserved factors into account.

The main parameter of interest is the coefficient $\gamma_{c',c}$ for all $c, c' \in C$, which expresses how a change in the semantic expression of i's friends affects subject i's semantic expression. However, the reciprocal causality present in model (1) makes it difficult to obtain unbiased estimates of the model parameters. This is due to the inherent feedback present in the model. That is, there is mutual influence between any pair of subjects i and j, and influence might follow even more complex paths (for example, i's expression in category c might influence j's expression in category c', which might affect k's expression in category c''). We address this issue in two steps, by first proposing an aggregated version of model (1) that averages over people who are in the same city (see Section III-B), and then by relying to the method of instrumental variable regression [28] (see Section III-C).

We also observe that model (1) is memoryless. This is a simplifying assumption that makes the method of instrumental variable regression easily applicable. Moreover, the model has one observation for each subject i and day $t \in T$, which, given a set of hundreds millions of users, would be difficult to analyze without some form of aggregation.

B. Geographical Aggregation

We average (1) over all $n_g(t)$ subjects $i \in S_g(t)$ who are in city g on day t

$$\frac{1}{n_g(t)} \sum_{i \in S_g(t)} y_i^{(c)}(t) = \frac{1}{n_g(t)} \times \sum_{j \in S_g(t)} \left(\theta(t) + f_i + \beta x_i(t) + \frac{\gamma}{\delta_i(t)} \sum_j a_{i,j}(t) y_i^{(c')}(t) + \epsilon_i(t) \right).$$

This can be written as

$$\bar{y}_g^{(c)}(t) = \theta(t) + \bar{f}_g + \beta \bar{x}_g(t) + \gamma \bar{Y}_g^{(c')}(t) + \bar{\epsilon}_g(t) \quad (2)$$

where we substituted $\gamma_{c',c}$ with γ and $\beta_{c',c}$ with β for ease of notation. In (2), $\bar{y}_g^{(c)}(t)$ is the average usage of category c by

[2]The model specification in (1) is not restricted to $c = c'$. It allows us to study the effect of usage of a semantic category c' on a potentially different category c.

subjects who are in city g; $\bar{f}_g$ is the average baseline usage of category c; $\bar{x}_g(t)$ is the average exogenous variable experienced by the subjects (rainfall in city g, in [21]); $\bar{\epsilon}_g(t)$ is the sum of $n_g(t)$ independent normal random variables with zero mean and variance σ^2 and, therefore, has variance $\sigma^2/n_g(t)$. The term $\bar{Y}_{gt}$ represents how usage of category c by subjects in city g is affected by the usage of category c' by their friends, and can be written as

$$\begin{aligned}\bar{Y}_g^{(c')}(t) &= \sum_j \frac{y_j^{(c')}(t)}{n_g(t)} \sum_{i\in S_g(t)} \frac{a_{i,j}(t)}{\delta_i(t)} \\ &= \sum_j y_j^{(c')}(t) A_{j,g}(t)\end{aligned}$$

where $A_{j,g}(t)$ represents the strength of the relationship from subject j to city g (normalized by the number of those subjects), that is, the influence from j to city g.

The coefficients β and γ are the same in (1) and (2). That is, the coefficients of the individual level model (1) can be estimated from the aggregated model (2). And note that our approach is unlikely to create an "ecological fallacy," which occurs when there are opposing effects at the individual and aggregated level, as individuals in the same city are very likely to experience the same weather [55]. Different instruments might lead to different situations.

Finally, the aggregated model (2) has a single observation for each city g and day t, a much smaller figure than the individual-level model (1), which would have millions of observations for each day in a large data set. This makes estimation more practical.

C. Instrumental Variable Regression

We are interested in estimating the parameter γ in (2). However, the explanatory variable $\bar{Y}_g^{(c')}(t)$ is an endogenous variable, that is, it can be correlated with both the dependent variable $\bar{y}_g^{(c)}(t)$ and the error term $\bar{\epsilon}_g(t)$. Since ordinary least squares regression would not produce unbiased estimates for γ, we use the method of instrumental variable regression [28]. This method can produce consistent and unbiased estimates even when there is reciprocal causation (as in our case, where people affect their friends and *vice versa*). All that is needed is an instrument that predicts the endogenous variable but not the dependent variable. More formally, given a linear model of the form

$$y = \alpha x + \lambda v + \epsilon$$

where v is an endogenous variable correlated with both the dependent variable y and the error term ϵ, an instrument for v is an exogenous variable z that does not appear in the model equation, is correlated with v (conditional on all the exogenous explanatory variables), and is not correlated with the error term [28]. Moreover, we look for a variable z such that, upon finding a relationship between z and v, z affects v and not *vice versa*. Once such variable z is available, instrumental variable regression estimates the original model in two stages. First, the endogenous variable v is projected onto the subspace of all exogenous explanatory variables, according to the model

$$v = \alpha_1 x + \alpha_2 z + \nu$$

where ν is an error term uncorrelated with any regressor. Then, the predicted values $\hat{v}$ resulting from the projection are used to estimate the model

$$y = \alpha x + \lambda \hat{v} + \epsilon.$$

In our model, an instrument for the endogenous explanatory variable $\bar{Y}_g^{(c')}(t)$ is an exogenous variable z that is uncorrelated with the error term in (2) [that is, $\mathrm{Cov}(z, \hat{\epsilon}_g(t)) = 0$] and is partially correlated with $\bar{Y}_g^{(c')}(t)$ when controlling for the other exogenous explanatory variables. In the context of our model, we can write

$$\bar{Y}_g^{(c')}(t) = \theta'(t) + \bar{f}'_g + \beta_2 \bar{x}_g(t) + \beta_1 z + \nu_g(t) \qquad (3)$$

where $\nu_g(t)$ is an error term that is uncorrelated with any regressors and $\theta'(t)$ and $\bar{f}'_g$ are time and subpopulation fixed effects.

Equation (3) can be seen as the linear projection of $\bar{Y}_g(t)$ on the space of all the exogenous variables. Substituting (3) into (2) yields

$$\begin{aligned}\bar{y}_g^{(c)}(t) = &(\theta(t) + \gamma\theta'(t)) + \left(\bar{f}_g + \gamma \bar{f}'_g\right) \\ &+ (\beta + \gamma\beta_2)\bar{x}_g(t) + \gamma\beta_1 z + \bar{\epsilon}'_g(t) \qquad (4)\end{aligned}$$

where the error term is uncorrelated with all the explanatory variables.

As the instrument z for $\bar{Y}_g(t)$, we define a variable $\bar{X}_g(t)$ that combines rainfall experienced by the friends of subjects in city g

$$\begin{aligned}\bar{X}_g(t) &= \sum_j x_j(t) \frac{1}{n_g(t)} \sum_{i\in S_g(t)} \frac{1}{\delta_i(t)} a_{i,j}(t) \\ &= \sum_j x_j(t) A_{j,g}(t) = \sum_h \bar{x}_h(t) \sum_{j\in S_h(t)} A_{j,g}(t) \\ &= \sum_h \bar{x}_h(t) B_{h,g}(t)\end{aligned}$$

where the sum is over all cities h, and

$$B_{h,g}(t) = \frac{1}{n_g(t)} \sum_{i \in S_g(t)} \frac{1}{\delta_i(t)} \sum_{j \in S_h(t)} a_{i,j}(t)$$

represents the strength of the relationship from city h to city g. We use $\bar{X}_g(t)$ to predict $\bar{Y}_g(t)$. $\bar{X}_g(t)$ is uncorrelated with the error term in (2), and it is partially correlated with $\bar{Y}_g^{(c')}(t)$.

The procedure above is equivalent to estimating the model in (2) using two-stage least squares (2SLS) regression. The first-stage regression estimates a model of the form

$$\bar{Y}_g^{(c')}(t) = \theta'(t) + \bar{f}'_g + \beta_1 \bar{X}_g(t) + \beta_2 \bar{x}_g(t) + \epsilon'_g(t). \quad (5)$$

The second-stage regression uses the predicted values $\bar{Y}_g^{(c',\mathrm{pred})}(t)$ from the first stage to estimate the model

$$\bar{y}_g^{(c)}(t) = \theta(t) + \bar{f}_g + \beta \bar{x}_g(t) + \gamma \bar{Y}_g^{(c',\mathrm{pred})}(t) + \bar{\epsilon}_g(t). \quad (6)$$

Finally, recall that the variance of the error term $\bar{\epsilon}_g(t)$ is proportional to $1/n_g(t)$ where $n_g(t)$ is the number of individuals in a city. Therefore, we weight each observation by the corresponding value of $n_g(t)$. To conduct the analysis, we use the function ivreg2 written for STATA [56].

D. Dealing With the Exclusion Restriction

A key assumption of instrumental variables regression is the exclusion restriction [28], according to which the instrument $\bar{X}_g(t)$ must not directly influence the dependent variable $\bar{y}_g^{(c)}(t)$. In our case, a person and some of her friends are experiencing similar $\bar{x}_g(t)$ as they are in the same city or in close-by cities. Therefore, in order to break the correlation between $\bar{X}_g(t)$ and $\bar{x}_g(t)$, we only consider observations for city-day pairs (g, t) such that $\bar{x}_g(t) = 0$ (in [21], it did not rain in city g on day t). Conditional on $\bar{x}_g(t) = 0$, (5) and (6) can be written as

$$\bar{Y}_g^{(c')}(t) = \theta'(t) + \bar{f}'_g + \beta_1 \bar{X}_g(t) + \epsilon'_g(t) \quad (7)$$

$$\bar{y}_g^{(c)}(t) = \theta(t) + \bar{f}_g + \gamma \bar{Y}_g^{(c',\mathrm{pred})}(t) + \bar{\epsilon}_g(t). \quad (8)$$

Note that since $\bar{x}_g(t) = 0$ the instrument $\bar{X}_g(t)$ now depends only on friends who are in different cities (not in city g). Therefore, our approach can only detect and measure influence between individuals in different cities.

E. Robustness of the Instrument

In order to assess the quality of the estimates obtained via instrumental variable regression, we also compute diagnostic statistics. First, we need to verify that the model is not underidentified. We use the Kleinbergen–Paap *rk* LM statistic to test the null hypothesis of underidentification [57]. Second, we need to verify that the instruments are good predictors of the endogenous explanatory variable in the first-stage regression (otherwise the instruments are considered weak). Weak instruments would cause poor predicted values in the first-stage regression and therefore poor estimation in the second-stage regression. To ensure the instruments are not weak, the Cragg–Donald Wald *F* statistic must exceed the critical threshold suggested by Stock and Yogo [58].

F. The Effect of a Person on Her Friends

We show that the coefficient γ represents the expected total effect of a person on her friends. In other words, it is the number of additional posts containing a word in category c posted by all of j's friends on day t caused by subject j's own post. Recall the individual-level model (1)

$$y_i^{(c)}(t) = \theta(t) + f_i + \beta x_i(t) + \gamma \frac{1}{\delta_i(t)} \sum_j a_{i,j}(t) y_j^{(c')}(t) + \epsilon_i(t). \quad (9)$$

Letting j be a subject who writes a post on day t, we compare the cases in which j's post contains a word in category c' ($y_j^{(c')}(t) = 1$) and that in which it does not ($y_j^{(c')}(t) = 0$). Simple manipulation of (9) shows that this difference is given by $\gamma a_{i,j}(t)/\delta_i(t)$. Summing over all subjects i who wrote a post on day t, the total effect of $y_j^{(c')}(t) = 1$ for a given subject j is

$$E_j(t) = \frac{\gamma \sum_i a_{i,j}(t)}{\delta_i(t)}. \quad (10)$$

The expected total effect of a person on all her friends is obtained by averaging (10) over all subjects j

$$\begin{aligned}\bar{E}(t) &= \frac{1}{n(t)} \sum_j E_j(t) = \gamma \frac{1}{n(t)} \sum_j \sum_i a_{i,j}(t)/\delta_i(t) \\ &= \gamma \frac{1}{n(t)} \sum_i \frac{1}{\delta_i(t)} \sum_j a_{i,j}(t) = \gamma \frac{1}{n(t)} \sum_i \frac{\delta_i(t)}{\delta_i(t)} = \gamma.\end{aligned}$$

Therefore, we can refer to the coefficient γ as the expected total effect of a person on her friends.

IV. RESULTS

In this section, we review the results from [21] to show how the method works. In future, we plan to apply the method to other semantic categories using data from a variety of social media platforms. The analysis in [21] was based on the posts written by a large sample of English-speaking Facebook users over a period of more than three years between 2009 and 2012, and we restricted our analysis to the categories of positive and negative emotions defined by the LIWC. Although these two categories are negatively correlated, they are not opposite sides of the same scale. Heightened emotional arousal might cause users to express themselves with both categories at the same time.

A. Model Parameters

Table 2 and Fig. 1(a) show that rainfall is a valid instrument for both categories of positive and negative emotion (reprinted from [21, Fig. 2A]). That is, it predicts enough of the variability of the content posted that it allows us to obtain reliable estimates of influence with our method. Table 3 and Fig. 1(b) show statistically significant estimates γ of contagion (reprinted from [21, Fig. 2B]). In particular, a person's post in one semantic category can cause friends to generate one to two additional posts in the same category (see Section III-F). Also, an increase in the usage of positive (resp., negative) emotion words by an individual inhibits the usage negative (resp., positive) emotion words by her social contacts.

B. Additional Tests

Since we would expect that friends' future expression does not predict a person's current semantic expression, we can consider the following placebo model:

$$\bar{y}_g^{(c)}(t) = \theta(t) + \bar{f}_g + \beta\bar{x}_g(t) + \gamma\bar{Y}_g^{(c')}(t+\delta) + \bar{\epsilon}_g(t) \quad (11)$$

where friends' future usage of category c' appears as an explanatory variable. We need to choose a lag of δ days in order to break the correlation between friends' present rainfall $\bar{X}_g(t)$ and future rainfall $\bar{X}_g(t+\delta)$. We can then estimate the model via 2SLS regression using friends' future rainfall $\bar{X}_g(t+\delta)$ as the instrument, and we would expect not to find statistically significant estimates of γ. In [21], we set $\delta = 30$ days and we found statistically insignificant estimates of γ for all considered models.

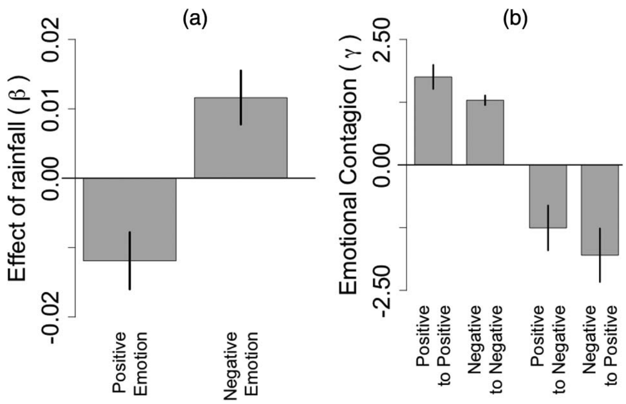


Fig. 1. ***(a) Effect β_1 of the instrument (friends' rainfall) on the endogenous explanatory variable (friends' positive and negative expression), from the first-stage regression. (b) Estimate of emotional contagion γ, from the second-stage regression. Vertical bars represent 95% confidence intervals. Reprinted from [21, Fig. 2].***

To test whether our estimates of influence are driven by people writing posts about the weather (a situation that would change our interpretation of the results), in [21], we considered a meteorological glossary supplied by the National Oceanic and Atmospheric Administration (NOAA),[3] and for each i and t, we defined $w_i(t)$ as the fraction of posts of subject i on day t containing a meteorological word. We consider the following version of model (1):

$$y_i^{(c)}(t) = \theta(t) + f_i + \beta_{c',c}x_i(t) + \delta w_i(t) + \gamma_{c',c}\frac{1}{\delta_i(t)}\sum_j a_{i,j}(t)y_j^{(c')}(t) + \epsilon_i(t)$$

and its aggregated version

$$\bar{y}_g^{(c)}(t) = \theta(t) + \bar{f}_g + \delta\bar{w}_g(t) + \beta\bar{x}_g(t) + \gamma\bar{Y}_g^{(c')}(t) + \bar{\epsilon}_g(t) \quad (12)$$

where $\bar{w}_g(t)$ is the average of $w_i(t)$ over all people in city g. The model is estimated via 2SLS regression, using $\bar{X}_g(t)$ as the instrument. Our results showed that when we control for weather-related words, the estimates of the influence coefficient γ for model (12) were unchanged with respect

Table 2 Estimates of the Coefficient β_1 (With Additional Statistics and 95% CI) for the First-Stage Regression of (7) for the Categories of Positive and Negative Emotion. *p* Values Smaller Than 0.05 Reject the Null Hypothesis of Zero Coefficient. The Kleibergen–Paap *rk* LM Statistics Reject the Null Hypothesis That the Regression Is Underidentified [57]. The Cragg–Donald Wald *F* Statistics Exceed the Critical Thresholds Suggested by Stock and Yogo [58] to Ensure the Instruments Are not Weak. All Statistics Are Robust to Heteroskedasticity, Autocorrelation, and Clustering. Reprinted From Tables 6 and 7 of the Supplemental Appendix to [21].

First-stage regression
Effect of the instrument $\bar{X}_g(t)$ on $\bar{Y}_g(t)$

Category	β_1	Standard Error	$P > \lvert t \rvert$	95% Conf. Interval Low	High
negative	0.0116	0.00195	0.000	0.00776	0.0155
positive	-0.0119	0.00207	0.000	-0.0160	-0.00781

[3]http://www.erh.noaa.gov/box/glossary.htm

Table 3 Estimates of the Coefficient γ (With Additional Statistics and 95% CI) for the Second-Stage Regression of Equation (8), When $c = c'$ and $c \neq c'$ for the Semantic Categories of Positive and Negative Emotion. p Values Smaller Than 0.05 Reject the Null Hypothesis of Zero Coefficient. Reprinted From Tables 5, 6, 7, and 8 of the Supplemental Appendix to [21]

Second-stage regression Influence by category					
Category	γ	Standard Error	$P > \|t\|$	95% Conf. Interval Low	High
Friends positive emotion (c')					
User positive emotion (c)	1.752	0.122	0.000	1.514	1.991
User negative emotion (c)	-1.255	0.227	0.000	-1.701	-0.809
Friends negative emotion (c')					
User positive emotion (c)	-1.798	0.271	0.000	-2.330	-1.266
User negative emotion (c)	1.288	0.0486	0.000	1.193	1.383

to those for the original model (2). This suggests that the influence estimate is not driven by people writing posts about the weather.

V. DISCUSSION

In [21], we proposed a rigorous method based on mathematical modeling and instrumental variable regression to detect and quantify contagion of semantic expression in online social networks using observational data. First, our method allows us to determine what semantic categories are susceptible to peer influence between social contacts. In particular, we showed that a person's post expressing positive or negative emotion can cause his or her friends to generate one to two additional posts expressing the same emotion. Second, it allows us to estimate a signed relationship between different categories, characterizing how an increase in the usage of a semantic category by an individual alters the usage of another by her social contacts. Third, our model allows us to quantify the cumulative effect that a person has on all her social contacts.

One potential concern is the instrument's weakness [42]; rainfall has only a small effect in our analysis, but this does not harm the validity of our conclusions because it is the precision, and not the size of the estimate, that matters. In the data set we used in [21], built from content posted by millions of users, even a small effect is statistically significant and robust to a multitude of statistical tests against instrument weakness.

Our method limits inference to influence between subpopulations (individuals in different cities). Drawing conclusions about influence within a subpopulation (individuals in the same city) using observational data requires either the identification of a valid instrument or the definition of a different approach. This is an avenue of future research.

There are, of course, some limitations in inferring causality from observational data, and robust instruments may not always be available. Our model provides an alternative method when a large scale experiment is infeasible and researchers must rely on observational data. In an experiment, one would directly control the state of some people in order to track changes in their friends' outcomes (semantic expression, in our case). With the proposed approach, which constitutes a "natural experiment," the instrument (rainfall, in our case) constitutes a source of variation that affects some people directly (those experiencing it) but can predict changes in their social contacts who do not directly experience it. Moreover, our method can be easily applied to massive data sets (thanks to aggregation), and allows us to perform multiple analyses regarding several outcomes.

We advocate for the involvement of the engineering community in the development of nonexperimental methods of causal inference. On the one hand, it is an open question how methods based on instrumental variable regression generalize to different contexts (especially contagion within a population) and how to build instruments in a systematic way. On the other hand, although instrumental variables might provide interesting answers, researchers should also develop and propose alternative techniques. ■

REFERENCES

[1] J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Disc. Data Mining*, 2009, pp. 497–506.

[2] S. Goel, D. J. Watts, and D. G. Goldstein, "The structure of online diffusion networks," in *Proc. 13th ACM Conf. Electron. Commerce*, 2012, pp. 623–638.

[3] N. A. Christakis and J. H. Fowler, "The spread of obesity in a large social network over 32 years," *New England J. Med.*, vol. 357, no. 4, pp. 370–379, 2007.

[4] N. A. Christakis and J. H. Fowler, "The collective dynamics of smoking in a large social network," *New England J. Med.*, vol. 358, no. 2, pp. 2249–2258, 2008.

[5] J. H. Fowler and N. A. Christakis, "Dynamic spread of happiness in a large social network: Longitudinal analysis of the Framingham Heart Study social network," *British Med. J.*, vol. 338, pp. 23–27, 2009.

[6] J. N. Rosenquist, J. H. Fowler, and N. A. Christakis, "Social network determinants of depression," *Molec. Psychiatry*, vol. 16, no. 3, pp. 273–281, 2011.

[7] S. Aral and D. Walker, "Identifying influential and susceptible members of social networks," *Science*, vol. 337, no. 6092, pp. 337–341, 2012.

[8] L. Muchnik, S. Aral, and S. J. Taylor, "Social influence bias: A randomized experiment," *Science*, vol. 341, no. 6146, pp. 647–651, 2013.

[9] S. Goel and D. G. Goldstein, "Predicting individual behavior with social networks," *Marketing Sci.*, vol. 33, no. 1, pp. 82–93, 2013.

[10] C. F. Manski, "Identification of endogenous social effects: The reflection problem," *Rev. Econ. Studies*, vol. 60, no. 3, pp. 531–542, 1993.

[11] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annu. Rev. Sociol.*, vol. 27, pp. 415–444, 2001.

[12] J. H. Fowler and N. A. Christakis, "Cooperative behavior cascades in human social networks," *Proc. Nat. Acad. Sci.*, vol. 107, no. 12, pp. 5334–5338, 2010.

[13] D. Centola, "The spread of behavior in an online social network experiment," *Science*, vol. 329, pp. 1194–1197, 2010.

[14] D. Centola, "An experimental study of homophily in the adoption of health behavior," *Science*, vol. 334, pp. 1269–1272, 2011.

[15] J. Guillory *et al.*, "Upset now? Emotion contagion in distributed groups," in *Proc. ACM SIGCHI Conf. Human Factors Comput. Syst.*, 2011, pp. 745–748.

[16] R. M. Bond *et al.*, "A 61-million-person experiment in social influence and political mobilization," *Nature*, vol. 489, no. 7415, pp. 295–298, 2012.

[17] A. D. I. Kramer, J. E. Guillory, and J. T. Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *Proc. Nat. Acad. Sci.*, vol. 111, no. 24, pp. 8788–8790.

[18] V. Goel, "Facebook tinkers with users emotions in news feed experiment, stirring outcry," *New York Times*, Jun. 29, 2014. [Online]. Available: http://www.nytimes.com/2014/06/30/technology/facebook-tinkers-with-users-emotions-in-news-feed-experiment-stirring-outcry.html

[19] M. Wood, "OkCupid plays with love in user experiments," *New York Times*, Jul. 28, 2014. [Online]. Available: http://www.nytimes.com/2014/07/29/technology/okcupid-publishes-findings-of-user-experiments.html

[20] M. Schroepfer, "Research at Facebook," Oct. 2, 2014. [Online]. Available: http://newsroom.fb.com/news/2014/10/research-at-facebook

[21] L. Coviello *et al.*, "Detecting emotional contagion in massive social networks," *PLoS ONE*, vol. 9, no. 3, DOI: 10.1371/journal.pone.0090315.

[22] J. W. Pennebaker, M. E. Francis, and R. J. Booth, "Linguistic inquiry and word count: LIWC." [Online]. Available: http://www.liwc.net

[23] J. W. Pennebaker, C. K. Chung, M. Ireland, A. Gonzales, and R. J. Booth, *The Development and Psychological Properties of LIWC 2007*. Austin, TX, USA: LIWC, 2007.

[24] Y. R. Tausczik and J. W. Pennebaker, "The psychological meaning of words: LIWC and computerized text analysis methods," *J. Lang. Social Psychol.*, vol. 29, no. 1, pp. 24–54, 2010.

[25] M. R. Mehl and J. W. Pennebaker, "The sounds of social life: A psychometric analysis of students' daily social environments and natural conversations," *J. Pers. Social Psychol.*, vol. 84, no. 4, pp. 857–870, 2003.

[26] J. W. Pennebaker, M. R. Mehl, and K. G. Niederhoffer, "Psychological aspects of natural language use: Our words, our selves," *Annu. Rev. Psychol.*, vol. 54, no. 1, pp. 547–577, 2003.

[27] J. Graham, J. Haidt, and B. A. Nosek, "Liberals and conservatives rely on different sets of moral foundations," *J. Pers. Social Psychol.*, vol. 96, no. 5, pp. 1029–1046, 2009.

[28] J. D. Angrist, G. W. Imbens, and D. B. Rubin, "Identification of causal effects using instrumental variables," *J. Amer. Stat. Assoc.*, vol. 91, no. 434, pp. 444–455, 1996.

[29] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts, "Everyone's an influencer: Quantifying influence on twitter," in *Proc. 4th ACM Int. Conf. Web Search Data Mining*, 2011, pp. 65–74.

[30] N. Agarwal, H. Liu, L. Tang, and P. S. Yu, "Identifying the influential bloggers in a community," in *Proc. ACM Int. Conf. Web Search Data Mining*, 2008, pp. 207–218.

[31] M. Cha, H. Haddadi, F. Benevenuto, and P. K. Gummadi, "Measuring user influence in Twitter: The million follower fallacy," in *Proc. 4th Int. AAAI Conf. Weblogs Social Media*, 2010, pp. 10–17.

[32] D. M. Romero, B. Meeder, and J. Kleinberg, "Differences in the mechanics of information diffusion across topics: Idioms, political hashtags, and complex contagion on Twitter," in *Proc. 20th ACM Int. Conf. World Wide Web*, 2011, pp. 695–704.

[33] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic, "The role of social networks in information diffusion," in *Proc. 21st ACM Int. Conf. World Wide Web*, 2012, pp. 519–528.

[34] S. Aral, "Commentary-identifying social influence: A comment on opinion leadership and social contagion in new product diffusion," *Marketing Sci.*, vol. 30, no. 2, pp. 217–223, 2011.

[35] S. Aral, L. Muchnik, and A. Sundararajan, "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks," *Proc. Nat. Acad. Sci.*, vol. 106, no. 51, pp. 21544–21549, 2009.

[36] J. D. Angrist and A. B. Krueger, "Empirical strategies in labor economics," *Handbook Labor Econ.*, vol. 3, pp. 1277–1366, 1999.

[37] D. Card, "The causal effect of education on earnings," *Handbook Labor Econ.*, vol. 3, pp. 1801–1863, 1999.

[38] G. M. Imbens and J. M. Wooldridge, "Recent developments in the econometrics of program evaluation," Nat. Bureau Econ. Res., no. w14251, 2008.

[39] J. R. Kling, J. B. Liebman, and L. F. Katz, "Experimental analysis of neighborhood effects," *Econometrica*, vol. 75, no. 1, pp. 83–119, 2007.

[40] J. Morduch, "The microfinance promise," *J. Econ. Literature*, vol. 37, no. 4, pp. 1569–1614, 1999.

[41] J. Bound, D. A. Jaeger, and R. M. Baker, "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak," *J. Amer. Stat. Assoc.*, vol. 90, no. 430, pp. 443–450, 1995.

[42] D. O. Staiger and J. H. Stock, "Instrumental variables regression with weak instruments," Nat. Bureau Econ. Res., no. 151, 1994.

[43] M. Stubbs, *Text and Corpus Analysis: Computer-Assisted Studies of Language and Culture*. Oxford, U.K.: Blackwell, 1996.

[44] M. Coulthard, Ed., *Advances in Written Text Analysis*. London, U.K.: Routledge, 2002.

[45] N. Fairclough, *Analysing Discourse: Textual Analysis for Social Research*. London, U.K.: Routledge, 2003.

[46] T. Joachims, *Learning to Classify Text Using Support Vector Machines: Methods, Theory and Algorithms*. Norwell, MA, USA: Kluwer, 2002.

[47] H. Lodhi, C. Saunders, J. Shawe-Taylor, N. Cristianini, and C. Watkins, "Text classification using string kernels," *J. Mach. Learn. Res.*, vol. 2, pp. 419–444, 2002.

[48] B. Sriram, D. Fuhry, E. Demir, H. Ferhatosmanoglu, and M. Demirbas, "Short text classification in twitter to improve information filtering," in *Proc. 33rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2010, pp. 841–842.

[49] X.-H. Phan, L.-M. Nguyen, and S. Horiguchi, "Learning to classify short and sparse text & web with hidden topics from large-scale data collections," in *Proc. 17th Int. ACM Conf. World Wide Web*, 2008, pp. 91–100.

[50] L. Jiang, M. Yu, M. Zhou, X. Liu, and T. Zhao, "Target-dependent Twitter sentiment classification," in *Proc. 49th Annu. Meeting Assoc. Comput. Linguist., Human Lang. Technol.*, 2011, pp. 151–160.

[51] R. Wardhaugh, *An Introduction to Sociolinguistics*. New York, NY, USA: Wiley-Blackwell, 2006.

[52] S. Dubois and B. Horvath, "Let's tink about dat: Interdental Fricatives in Cajun English," *Lang. Variat. Change*, vol. 10, no. 3, pp. 245–261, 1998.

[53] E. Gilbert and K. Karahalios, "Predicting tie strength with social media," in *Proc. ACM SIGCHI Conf. Human Factors Comput. Syst.*, 2009, pp. 211–220.

[54] R. Xiang, L. Neville, and M. Rogati, "Modeling relationship strength in online social networks," in *Proc. 19th ACM Int. Conf. World Wide Web*, 2010, pp. 981–990.

[55] W. S. Robinson, "Ecological correlations and the behavior of individuals," *Amer. Sociol. Rev.*, vol. 15, no. 3, pp. 351–357, 1950.

[56] C. F. Baum, M. E. Schaffer, and S. Stillman, "Enhanced routines for instrumental variables/GMM estimation and testing," *Stata J.*, vol. 7, no. 4, pp. 465–506, 2007.

[57] F. Kleibergen and R. Paap, "Generalized reduced rank tests using the singular-value decomposition," *J. Econometrics*, vol. 127, pp. 97–126, 2006.

[58] J. H. Stock and M. Yogo, "Testing for weak instruments in linear IV regression," in *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, D. W. K. Andrews and J. H. Stock, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2005.

ABOUT THE AUTHORS

Lorenzo Coviello (Student Member, IEEE) received the Laurea degree (*magna cum laude*) in telecommunication engineering from the University of Padova, Padova, Italy, in 2008 and the M.S. degree in electrical and computer engineering from the University of California San Diego, San Diego, CA, USA, in 2013, where he is currently working toward the Ph.D. degree in electrical and computer engineering, under the advice of Prof. M. Franceschetti.

His research is focused on social networks, big data, and algorithms.

James H. Fowler received the B.A. degree (*cum laude*) from Harvard University, Cambridge, MA, USA, in 1992, the M.A. degree in international relations from Yale University, New Haven, CT, USA, in 1997, and the M.A. and Ph.D. degrees in government from Harvard University in 2001 and 2003, respectively.

He is a Professor of the Department of Political Science and School of Medicine, Medical Genetics Division, San Diego, CA, USA. His work lies at the intersection of the natural and social sciences, with a focus on social networks, behavior, evolution, politics, genetics, and big data.

Massimo Franceschetti (Senior Member, IEEE) received the Laurea degree (*magna cum laude*) in computer engineering from the University of Naples, Naples, Italy, in 1997 and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology (Caltech), Pasadena, CA, USA, in 1999, and 2003, respectively.

He is a Professor in the Department of Electrical and Computer Engineering, University of California San Diego, San Diego, CA, USA. He was a Postdoctoral Scholar at the University of California Berkeley, Berkeley, CA, USA. He has held visiting positions at the Vrije Universiteit Amsterdam, Amsterdam, The Netherlands; the Ecole Polytechnique Federale de Lausanne, Lausanne, Switzerland; and the University of Trento, Trento, Italy. His research interests are in communication systems theory and include random networks, wave propagation in random media, wireless communication, and control over networks.

Dr. Franceschetti is an Associate Editor for Communication Networks of the IEEE Transaction on Information Theory (2009–2012) and has served as Guest Editor for two issues of the IEEE Journal on Selected Areas in Communication. He was awarded the C. H. Wilts Prize in 2003 for best doctoral thesis in electrical engineering at Caltech; the S.A. Schelkunoff Award in 2005 for best paper in the IEEE Transaction on Antennas and Propagation; a National Science Foundation (NSF) CAREER award in 2006; an Office of Naval Research (ONR) Young Investigator Award in 2007; and the IEEE Communications Society Best Tutorial Paper Award in 2010.

INVITED PAPER

Analyzing Temporal Networks in Social Media

This paper reviews methods for analyzing recently temporal networks applied to social media data.

By Petter Holme

ABSTRACT | Many types of social media metadata come in forms of temporal networks, networks where we have information about not only who is in contact with whom but also when contacts happen. In this paper, we review methods to analyze temporal networks developed in the last few years applied to social media data. These methods seek to identify important spreaders and, in more generality, how the temporal and topological structure of interaction affects spreading processes.

KEYWORDS | Network analysis; social network services; temporal networks

I. INTRODUCTION

Many types of interactions in social media form temporal networks—collections of unique interaction events between pairs of individuals [1]. The structure of these temporal networks determines the dynamics of information spreading. To study such phenomena is of interest, not only from an academic standpoint, but also for word-of-mouth marketing and similar types of technologies that rely on information spreading between people.

To understand how a large, integrated system functions as a whole, one needs to zoom out and look at it from a distance. In other words, one needs a consistent way of simplifying and discarding irrelevant information. A solution that has gained a huge amount of interest during the last decade goes under the names of complex networks, or network theory. In this paradigm, one only keeps the information about the interacting units (in social media that would typically be individuals or advertisers) and who has been in contact with whom. Usually, even if one is concerned only with metadata and not message content, one has more information than only this static network. The advantage with looking only at a static network (a simple graph in mathematics jargon) is that one has a huge toolbox for analyzing the data. If one also includes information about when interaction events occur, i.e., studies a temporal network rather than a static network, then the number of methods is much more restricted. The good news is that the study of temporal networks has been a very active field in the last few years, so the number of methods that study synergetic temporal and topological effects is increasing fast.

In this paper, we will review temporal network methods and discuss how they can be used to understand spreading events, identify influential spreaders, and describe social organization as reflected in activities on social media.

II. TEMPORAL NETWORKS AS A MODELING FRAMEWORK

A. Representations and Types of Interactions

The two main classes of mathematical representations of temporal networks are contact sequences and interval graphs. These two concepts are more similar than their names suggest. In contact sequences, a contact is assumed to be instantaneous with respect to the temporal granularity of the data. They could thus be represented as triples (i, j, t) encoding an interaction between individual i and j at time t. In interval graphs, one can picture an interaction event as being temporally extended, so it has a beginning and end time. Mathematically, a contact would be represented as a quadruplet rather than triple (i, j, t, t'),

Manuscript received April 25, 2014; revised August 14, 2014; accepted September 29, 2014. Date of publication November 4, 2014; date of current version November 18, 2014. This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education under Grant 2013R1A1A2011947.
The author is with the Department of Energy Science, Sungkyunkwan University, Suwon 440-746, Korea and also with IceLab, Department of Physics, Umeå University, Umeå 90187, Sweden (e-mail: holme@skku.edu).

Digital Object Identifier: 10.1109/JPROC.2014.2361326

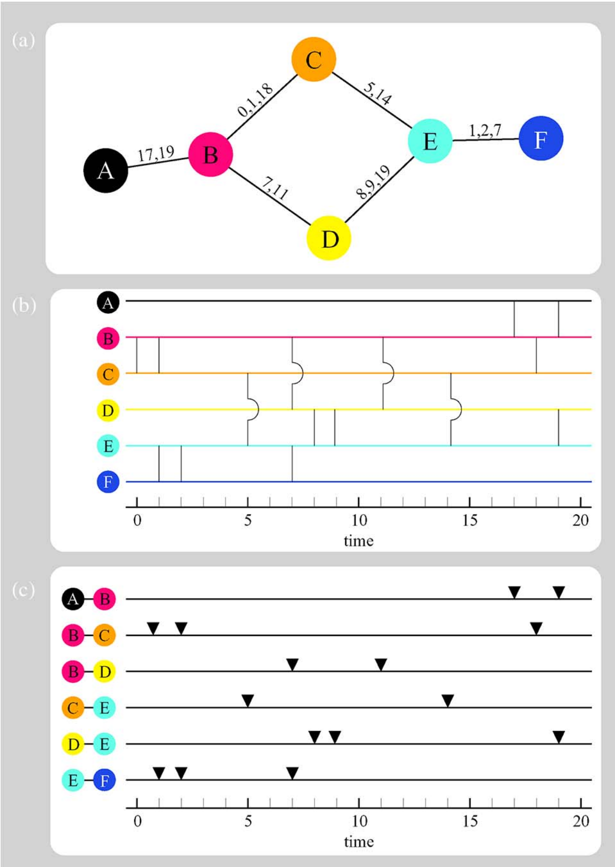


Fig. 1. ***Graphical representations of temporal networks. This figure shows three ways of displaying a temporal network (more specifically a contact sequence). Panel A shows a plot of the network of aggregated contacts with the edges annotated with the times of contacts. Panel B shows a timeline of the vertices where a contact is indicated by a horizontal line. Panel C displays a timeline of edges where each contact between the vertex pair is indicated by a triangle. This temporal network illustrates the nontransitive nature of temporal networks. Information spreading over the contacts could reach from A to B (e.g., via the contact at time 17) and from B to D, but not from A to D via B, because by the time the information would have reached B, all contacts between B and D have already happened.***

meaning that i has a contact with j that lasts from time t to t'. In social media, the interaction is usually directed, so that the order of the individuals i and j carries a meaning in the contact representations. In many cases, the same methods can be used for contact sequences and interval graphs. The latter can always be reduced to the former at the expense of some information loss. In this paper, we focus on contact sequences. Fig. 1 illustrates some ways to draw a contact sequence that displays all its information. It is a system of six vertices, and already quite complicated. Unfortunately, displaying social media data such as these figures would not be very informative. They serve as a base for the discussion of our concepts.

Social media interaction that would suit the contact sequence format includes direct (e-mails or e-mail like) messages [2], [3], Facebook wall posts [4] and friendships [5], [6], Internet forum posts [7], [8], blogs [9], microblogging [10], location sharing [11], etc. Usually, assigning a time of an event is not completely trivial. From a spreading point of view, the relevant time would be when the recipient is exposed to the information, but it is often the sending time that is stored in the logs of the media. Most studies use this reception time as the only time stamp of the contacts. This causes the short time order of events to be jumbled and alters the paths that the information can flow through, but probably not changes the results for larger spreading events. (A thorough study of this bias would be interesting.)

Interaction that takes the form of a dialog, either over a voice or chat channel could be better modeled as an interval graph. This type of data is rather well studied when it comes to mobile phone communication [12]–[14] (perhaps a border case of social media). We are not aware of any study addressing voice call features embedded in social media platforms (perhaps it makes sense not to classify these as social media anyway). In the rest of this paper, we will restrict our discussion to contact sequences. One reason is that one can, and often does, reduce interval graphs to contact sequences, either by discretizing time and logging one contact in every interval of the contact, or by taking the beginning (or the middle) of the interval as the interaction time.

B. Models for Spreading Processes

There is a large and fast evolving field studying social influence and other forms of information spreading between people that are applicable to social media. There are two main classes of such models: simple and complex contagion models. The first class seeks to capture processes where only one individual affects another (relays a rumor, changes an opinion, etc.). Input from more than one person is negligible with respect to the system-wide dynamics. There is a popular idea that information (Internet memes as a notable example [15]) spreads like infections (as reflected in expressions such as "go viral"). Also, in the academic literature, the idea goes back about half a century [16]. Even if this analogy can capture the fundamental features of spreading, it can probably not hold in a strict sense: transmission probabilities in information spreading probably have much more individual variation than infectious disease spreading, and they can be related to ongoing events (elections in the examples of [15]) which is something one would never see in disease transmission. Nevertheless, if the fundamental properties of contagion are the same as for infectious diseases, then one can, of course, borrow models from epidemiology to understand spreading in social media. The basic class of models for infectious disease spreading is called compartmental models [17]. These models divide people into classes, or compartments, with respect to the disease: canonical classes include susceptible (S), infectious (I), and recovered (R). The latter class comprises individuals that have stopped being infectious and acquired immunity to the pathogen, or died. In a social media context, R would correspond to individuals that have had the information but stopped being interested in spreading it further (like a once viral video that people have lost interest in). The second defining feature of compartmental models is that they assign transition rules between the classes. One of the rules that is (to the best of our knowledge) always included is that a susceptible, upon encountering an infectious, with some probability can become infectious.

Another type of simple contagion models are the ones derived from the voter model [18], [19]. These are typically nonequilibrium models to show spreading of an opinion within a population. In the standard setting, one lets the opinion of a person be an integer valued variable. Then, one updates the system by iteratively picking a random node i, then picking a random neighbor j of i and finally copying j's opinion to i. This model is analytically tractable, but other than that it may be something of an oversimplification as a model of real information spreading. Typically, one is interested in the time it takes for the voter model to reach a state where every node has the same opinion as its neighbors.

In complex contagion, spreading can be contingent on the interaction with more than one other person. A popular idea is that opinion spreads following a threshold model [20]. The idea is that social contagion could happen if someone has been influenced by a certain fraction of others. In a temporal network, one would also have to decide how to take time into account. For most cases, social influence that happened a long time ago will not matter in, for example, the adoption of a product or spreading of an opinion. Karimi and Holme [21] discuss this further and use a sliding time window to represent the duration of possible influence. Takaguchi *et al.* [22] use an exponentially decaying weight for the same purpose in a bit more elaborate model of social contagion.

C. Time-Respecting Paths, Components, and Distances

Any process that is confined to a temporal network has to follow sequences of contacts of increasing times. (In some situations, depending on the spreading process in question, one would also allow nondecreasing times, so that the spreading could pass two contacts at the same time step.) There are a few different names for such a sequence of contacts; we will follow Kempe and call it a time-respecting path [23]. Although a time-respecting path, as a concept, is the counterpart of the paths in static graphs, it has some fundamentally different properties. To begin with, if, at time t, one can reach node j from node i by a time-respecting path, this does not imply that there is a time-respecting path from j to i at time t. In other words, the property of i being connected to j is not commutative. This property is the same for directed graphs. As a consequence, just like for directed networks, one can talk about strongly and weakly connected components in temporal networks. A strongly connected component is a set of vertices where one can reach from any vertex to any other vertex following time-respecting paths (e.g., B, C, D, and E in Fig. 1). To define a weakly connected component, we first construct an undirected graph where edges represent pairs that have at least one contact throughout the sampling time. A connected component in this graph would then be a weakly connected component in the original temporal network.

There are, however, other features of temporal networks that set them aside from directed graphs. Most conspicuously, time-respecting paths are not transitive. This means that even if, at time t, there are time-respecting paths from i to j and from j to k, there are not necessarily any time-respecting paths from i to k. This is the case if all time-respecting paths from j to k have already happened by the earliest time that a time-respecting path from i reaches j. As a word of caution, since a time-respecting path is a collection of contacts that happen at certain times, a statement such as "there is a time-respecting path between i and j" is ambiguous. One always needs to specify a time t, meaning that at t you can reach j from i following a time-respecting path. In Fig. 1, there is a time-respecting path from F to A at time 5 but not at time 10. Also, for components, one needs to specify a time. An interesting extension of the component concept would be to consider sets of vertices that are, within a time window, transitive.

A useful concept related to time-respecting paths is the set of vertices that, at a certain time, can be reached by time-respecting paths from vertex i, which is called the set of influence of i. This is important for spreading processes, as it is the set of vertices that can eventually be influenced by i. Some studies have measured the reachability ratio: the average fraction of vertices in sets of influence averaged over all beginning times during the sampling time [24]. Similarly to the set of influence, one can also define the source set of i: the set of vertices that can reach i through time-respecting paths within an observation window. This set consists of all vertices that can be the source of a spreading process influencing i. The work of Moody [25] is the earliest we are aware of that studies the size of the source set (the source count).

Of course, since the source set and set of influence are time specific, one may also monitor the reachability ratio and source count functions of time, i.e., study how many other vertices may reach vertex i by time-respecting paths by time t', when the paths begin no earlier than $t < t'$.

For static graphs, the distance between two vertices is defined as the length of the shortest path joining them (where path length is defined as the number of links forming a path). Short average distance is, of course, assumed to be a sign that a temporal network is efficient with respect to spreading. One of the main findings for static networks is that they are often remarkably compact. This "small-world phenomenon" exists in virtually all empirical networks except where links are physical objects (such as roads, powerlines, etc.). Naturally, when the dimension of time is added to the picture, it is useful to define similar quantities characterizing how quickly vertices can reach each other through time-respecting paths. Cooke and Halsey studied this kind of quantity in the 1960s [26]. The precise definitions of such quantities are not completely straightforward. How should one, for example, compare the two, if one exists half of the sampling time and another only 10%.

A time-respecting path is associated with duration, measured as the time difference between the last and first contacts on the path. (Even though the dimension is time, some authors have called it the temporal path length [12].) So the duration of the path from F to A in Fig. 1 is 18 (the first contact that can be involved in spreading is between F and E at time 1; the last contact is between B and A at time 19). Analogously to the shortest paths of static graphs, which define the distance, one can study the fastest time-respecting path(s) between two nodes.

The concept of latency was originally introduced in the field of distributed computation [28]. A central problem in this area is to keep track of the age of information that a vertex has about other vertices. Then, one commonly assumes that vertices in contact update each other's information so that, after contact, both vertices share the most recent information that either of them had before contact. This scenario is similar to the fastest spreading limit of spreading processes (for example, a disease-spreading model with 100% per-contact infection probability).

We will sketch the framework of latency as introduced by Lamport [27] and further developed by Mattern [28]. Consider information spreading in a temporal network and, specifically, vertex i. Then, let $\phi_{i,t}(j)$ denote the latest time such that information from j could have reached i by time t. This quantity is called i's view of j's information at time t. Furthermore, $\lambda_{i,t}(j) = t - \phi_{i,t}(j)$ is called j's information latency, or just latency, with respect to i at time t, and is thus a measure of how old i's information coming from j is at time t. For example, the latency from F to A in Fig. 1 at $t = 17$ is 10. In other words, $\phi_{A,17}(F) = 7$, so $\phi_{A,17}(F) = 17 - 7 = 10$. Finally, vector $[\phi_{i,t}(1), \ldots, \phi_{i,t}(N)]$ is called the i's vector clock. A difference to other approaches is that this concept is looking backward in time. Looking forward in time, one may define a quantity corresponding to latency (temporal distance [12]) $\tau_{i,t}(j)$ that measures how long it takes to reach j from i along the fastest path, starting the clock at time t. The expected temporal distance for a random starting time is called reachability time [24]. An adaptation of the vector-clock concept to social media data can be found in [29].

As we allude to above, latency and vector clocks form a basis for measuring times and optimal spreading speeds in temporal networks. However, taking an average over the sampling time to get a value for the entire graph, or even only for a pair of vertices, is not that straightforward. Problems are typically related to the finite time windows of empirical data sets. For example, as the time gets closer to the end of the sampling, the number of time-respecting paths between a pair of vertices decreases. If the sampling would be longer, one could presumably see more time-respecting paths starting even before the last time-respecting paths. One possible quantity for measuring the velocity of paths in general is to find all fastest time-respecting paths between vertices and then compute the average duration of such paths. This measure would, however, not reflect the frequency of the paths, and would not be affected by waiting times before the first contacts of such paths. For example, if one or ten time-respecting paths of one unit connect two vertices, this average duration would equal unity in both cases.

Measuring the average latency is also complicated by the fact that latency varies with time with a sawtooth pattern [1]. The latency goes up linearly until there is contact that is the beginning of a new time-respecting path from the source node j, carrying newer information from i than i already has. Close to the beginning of the observation window, the latency is infinite as no time-respecting path has reached its destination yet. In a steady-state situation where all links are active throughout the sampling time, it would be a good approximation to assume that the proceeding time interval of the sampling duration is similar to the first. This suggests a boundary condition where one repeats the entire temporal contact sequence, and thereby gets around the problem where there are few paths in the beginning of the sampling time. However, this procedure may give rise to artifacts and connect pairs of vertices that are not connected at all within the observation window. Yet another option would be to average the time between the first and last contacts, which would underestimate any spreading processes, especially for short paths.

For long enough periods of observation, another difficulty comes from the dynamics of vertices entering and leaving the system. For disease spreading in empirical data sets, this turnover of vertices and edges has been argued to be of great importance [30], [31]. In such a case, the question about how large a spreading event can be is probably more important than the time it takes to reach between the vertices.

Finally, here are some words of caution about the terminology. Some authors use the terms as distance and length as measures of time, e.g., Kossinets *et al.* [32] define the "distance" between two vertices as the shortest duration of any time-respecting path between them. Tang *et al.* [33] calls the average time to reach vertices for time-respecting paths starting early in the data

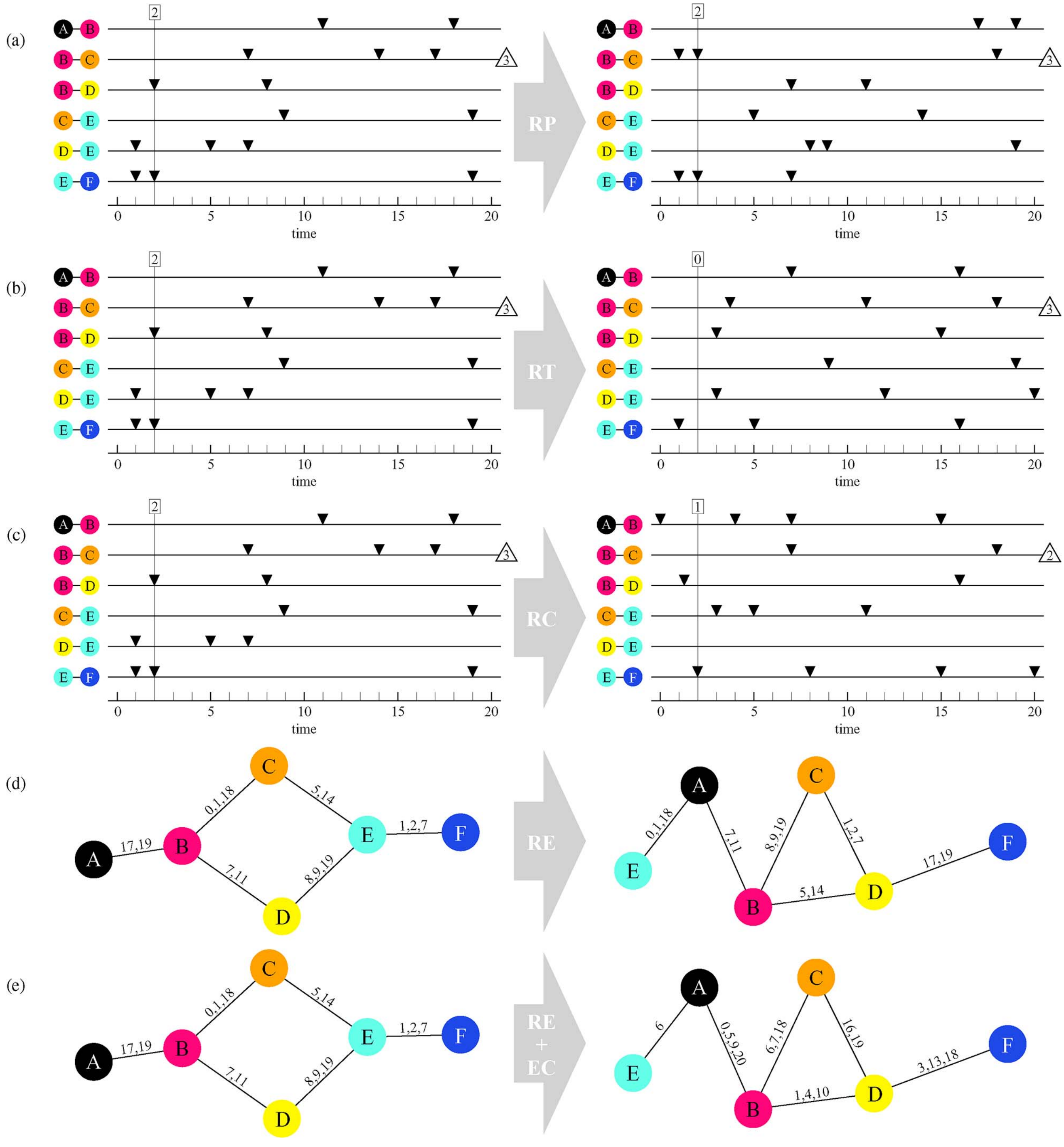


Fig. 2. ***Illustration of various randomization schemes. Each panel shows a possible instance of a randomization scheme applied to the network in Fig. 1. In panel A, the random permuted (RP) times method, the time stamps are randomly shuffled between the contacts. The resulting temporal network has the number of contacts per time and the number of contacts per edge conserved. For example, at time 2, there are two contacts both before and after applying the randomization scheme (see the boxed number). Similarly, the edge (B, C) has three contacts before and after randomization. In panel B, we show the random time (RT) scheme where time stamps are not swapped (as in RP) but created. Thus, this randomization does not conserve the number of contacts per time period (so, as illustrated, at time 2 the number of contacts changes from two to zero). In C, we illustrate the random contact (RC) randomization, which does not conserve the number of contacts per edge. After randomization, the (B, C) edge has two contacts, rather than three. Panel D shows the randomized edge (RE) scheme that keeps the degrees in the network of accumulated contacts constant, but randomizes the edges. Furthermore, the timelines of edges are held fixed. Panel E shows the combination of RC and RE where the only thing conserved is the degrees in the***

"temporal path length." To be fair, we should point out that "distance," as in the standard static graph definition, is also a misnomer, being a dimensionless quantity. Furthermore, statistics of latency have several names even though their definitions and purpose are very similar: "reachability time" in [24], "temporal proximity" in [34], "characteristic temporal path length" in [33], and "temporal distance" in [12].

D. Randomization Techniques

In static networks, network structures are a key to understanding both the forces that shape the network and how dynamic processes on the network behave. The loose definition of network structure is that it is the way the network differs from a random network. To make this definition operational, one commonly specifies a null model, a model for how to construct a network that is random except the most fundamental constraints. What is fundamental depends on the system one studies and the type of analysis one wants to do. Probably the most common null model for static networks are networks with the same set of degrees (and consequently the same number of nodes and links) as the original network. One can easily sample such networks by randomizing edges: one picks random pairs of edges (i, j) and (i', j') and swaps them to (i, j') and (i', j) or (i, i') and (j, j') (unless swapping would introduce a self-link or a multiple-link). Then, one could compare quantities, such as the number of triangles, in the real network to the average number of triangles in the null model. If there is an overrepresentation of triangles in the real network compared to the null model (which is commonly the case in social networks), one can conclude that the network was formed by a process that has a bias for triangle formation (such as people getting acquainted with one another by being introduced by a common friend). Furthermore, one can study processes on the network (spreading phenomena, for example) and compare properties of the process in the real networks and the randomized networks.

Comparing the real data with a null model based on randomizing the real data becomes even more powerful in temporal networks; it is probably the only way to consistently analyze temporal and topological features of the data, and thus to discover synergetic features where time and topology together influence a spreading process. Since there are several possible temporal correlations and several time scales where the correlations can affect the system, there will not be any method as common as the random link shuffling procedure described above for static networks. Rather, by designing appropriate null models, one may switch off different types of correlations in order to understand their contribution to some quantity describing a process on the temporal network (such as a measure of the spreading speed). A typical use for such models in studies of spreading processes would be to apply them sequentially, and by monitoring how the dynamics of the process depends on these null models, to pinpoint the role of different temporal and topological correlations on the process. If removing a certain type of correlations changes the dynamics more than another, then obviously the first played a more important role for the spreading process.

Below we review temporal-network null models introduced in the literature (some of them are illustrated in Fig. 2). The first paper using this type of methods was (to the best of our knowledge) [24]. At the end of the section, we summarize and provide some guidelines for choosing reference models. A summary of the randomization techniques can be found in Table 1.

1) Randomized Edges (REs): This method is similar to the edge swapping for static graphs mentioned above, with the additional fact that contact sequences of edges follow the edges when these are rewired. In pseudocode, the method is defined as follows.

1) Go over all edges sequentially.
2) For every edge (i, j), pick another edge (i', j').
3) With a probability 1/2, replace (i, j) and (i', j') by (i, j') and (i', j); otherwise, replace them by (i, i') and (j, j').
4) If the move in step 3) created a self-edge or a multiple-edge, then undo it and start over from step 1).

The times of contact over an edge are kept constant. Note that the two alternatives in step 3) where one is randomly selected are needed to remove spurious

Table 1 Summary of the Randomization Techniques and Which Structures They Preserve (Everything Else Is Randomized)

Randomization method	Conserved structures
Randomized edges	Time sequence of edges. Degree distribution.
Randomly permuted times	Network of accumulated contacts. Number of contacts per edge. Times of contacts in general.
Random times	Network of accumulated contacts. Distribution of number of contacts per edge.
Randomized contracts	Network of accumulated contacts. Times of contacts in general.
Equal-weight edge randomization	Network of accumulated contacts. Number of contacts per edge.
Edge randomization	Network of accumulated contacts. Distribution of number of contacts per edge. Times of contacts in general.
Time reversal	Everything except the direction of time.

correlations if (as is usually the case) the data structure store edge returns the vertices of an edge in a specific order. Without this rule, one would keep the number of times a vertex appears in the first argument conserved, which could give very different results. This algorithm is usually not a bottleneck in a temporal-network analysis, but to speed up the process, one can skip edges that already have been rewired in 3).

Just as for static network, this null model can be used to study the effects of the network topology (apart from the degrees, the number of neighbors). The model implicitly also assumes that it is the edges rather than the vertices that govern the times of contacts. This is evident since, after the randomization procedure, both the number and times of contacts for a vertex will probably have changed. However, their degrees in the aggregated network are the same. As the contact sequences follow their edges when rewiring, all temporal correlations and inhomogeneities associated with individual edges, such as burstiness and the distribution of intercontact times of edges, are retained. Also the system-wide contact rates are unaltered.

2) Randomly Permuted (RP) Times: To understand the role of the order of the contacts, one performs the RP randomization. In this procedure (used in, e.g., [14] and [35]), one permutes the contact times randomly while keeping the network structure and the numbers of contacts between all pairs of vertices fixed. Technically, this is much simpler than applying the edge-rewiring scheme discussed above, as it only requires randomly swapping the time stamps of all contacts. No checks similar to step 4) of the RE rule need to be performed for contact sequences. Like RE, this scheme also retains the overall rate of events in the network at every point in time, such as daily or weekly patterns in communication networks.

3) Random Times (RTs): The ensemble defined by RE and RP randomizations conserves the set of times of the original contact sequence. Hence, although it destroys time structures of events related to individual vertices and edges, the rate of events in the entire temporal network is unchanged and will still follow the typical circadian and weekly patterns of human activity (see, e.g., [36]–[39]). This type of randomization needs a process where spreading depends on time, not only the order of contacts (such as the SIR model of disease spreading, where individuals would change state even though no contact happens). In maximal speed spreading such as the one behind the latency discussion above, this randomization would have exactly the same effect as RP. Note that an alternative to uniformly random contact times is to generate them from a specific distribution or process, such as the Poisson process, with parameters set up so that the numbers of contact per each edge are, on average, conserved.

4) Randomized Contacts (RCs): For this randomization scheme, one keeps the graph topology fixed but redistributes the contacts randomly among the edges. After this randomization, the number of contacts per edge follows the binomial distribution. It is intended to test the effect of fat-tailed distributions of this quantity as typically seen, especially in social media and other forms of human-generated communication data. If one would like to test the effect of the distribution of the number of contacts alone, keeping the structure of the temporal order of the real data, then one would need a different approach. For example, a vertex that is active primarily in the early stage of the data would be so in the randomized data as well, but one would need to compensate for such effects.

5) Equal-Weight Edge Randomization (EWER): Sometimes one would need to remove correlations between the static network structure (the network of accumulated contacts), and at the same time retain the temporal structure of the edges (including the interevent time distributions of edges). This is achieved by randomly swapping entire contact sequences of edges with the same number of contacts. For example, all the contacts and their time stamps are randomly exchanged between edges that have the same number of contacts. Thus, single-edge patterns, such as burstiness of the contacts between two individuals, are retained, together with other properties preserved by the RP model (like the number of contacts of an edge, the system-wide contact frequency, and the topology of the network of accumulated contacts). This null model requires a large enough system so that there are enough edges with the same number of events.

6) Edge Randomization (ER): This null model is similar to the EWER model with the exception that the sequences can be exchanged between edges that have any numbers of contacts. This corresponds to randomly exchanging the edge weights (measured as numbers of contacts) in the network of aggregated contacts. The correlation between weight and topology is destroyed in this null model. However, the intercontact time distributions of contact sequences of edges are preserved; the sequences are just moved elsewhere in the network. Both EWER and ER were introduced in [12].

7) Time Reversal (TR): This null model is designed for assessing the frequency and importance of causal sequences [40] of contacts, where, for example, contacts trigger further contacts. It simply involves running the original event sequence backward in time. If sequences of consecutive contacts would be caused by temporal correlations alone, similar numbers of such sequences should be observed when time runs forward and backward. A lack of such chains in the time-reversed null model compared to the original sequence could be attributed to the arrow of time.

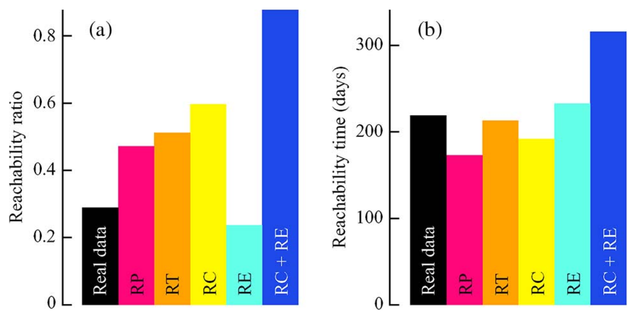


Fig. 3. ***Reachability ratio and reachability time analysis for interactions on an Internet dating site. Panel A shows the reachability ratio for the empirical temporal network: the probability there is a time respecting path between a pair of vertices at a random time. Panel B shows the reachability time: given there is a time-respecting path from one vertex to another, then what is the expected time for the information to complete this path (assuming it spreads over a contact whenever it can)? Averages over randomizations are over 104 runs. Standard errors are not larger than 0.01 in A and 0.5 in B (and thus too small to be meaningful as error bars).***

8) Summary and Guidelines: The different randomized reference models discussed above retain and destroy specific kinds of topological and temporal correlations, for example, in studies of processes on the temporal network. They allow for pointing out the importance of various correlations; the most important correlations can be pinpointed by comparing the effects of different randomization models on the dynamics. The RE and RC models permute edges and contact times. Taken together (RC + RE) they destroy all correlations except for the degrees of the network of accumulated contacts; this provides a good starting point for the limiting case of uncorrelated temporal networks. If one randomizes the contact times (RT), the overall contact frequencies are also removed. When studying the roles of the exact contact timings on edges and the correlations between adjacent edges, comparing the EWER and ER models to the RT model should work, as the static network features are retained, except for correlations between weight and topology that can be removed by ER. An illustration of some of these randomization schemes can be found in Fig. 2.

In Fig. 3(a), we show an example of these techniques applied to a social media data set. The temporal network in this case records the interaction (private messages, posts on home pages, and friendship requests) on an Internet dating site [41]. The data set records 536 276 contacts over 174 662 edges between 29 341 users for 512 days. First, we note that there is a very strong effect of the temporal ordering. These randomization schemes all keep the network of accumulated contacts unchanged. Still the reachability ratio can differ more than a factor of three. Even a very mild randomization such as RP, which tests the effect of the order of an event, increases the reachability ratio from 29% to 47%. Comparing RP and RT, the reachability ratio increases further, meaning that the actual time things happen makes spreading reaching fewer people. Applying the RC scheme makes spreading yet less efficient, so the heterogeneous distribution of contacts per edge [24] also limits the spreading. Randomizing the edges, on the other hand, decreases the reachability ratio from its original value, but not much. Changing the temporal properties (such as the previously mentioned randomization schemes) is thus more influential than the topology of the network of accumulated contacts. Further proving this point, the combination of RE and RC gives the largest reachability ratio. In Fig. 3(b), we show reachability times for the same data set as in panel A. All the randomizations that preserve the network topology speed up spreading, meaning that all (considered) temporal structures slow down spreading. This has been called the "slow-world effect" [14] (paraphrasing the "small-world effect," empirical static networks do often have very short average path lengths). The slow-world effect is not completely universal. Rocha *et al.* [35] find the reversed situation in a network of sexual contacts reported on a web forum.

E. Centrality and Influential Spreaders

There has been a tremendous interest in identifying important spreaders in the social media literature (see, e.g., [42]–[44] and further references therein). Note that "influential" (someone who are likely to influence many others) is not necessarily the same as "important" (someone who is able to boost a spreading process). Important spreaders also need to be susceptible to influence [42]. Except obvious applications such as seeding word-of-mouth advertisement [42]–[44], finding important spreaders could, for example, be useful in the detection of disease outbreaks and other applications in public health (this is discussed further in Section III).

In static networks, concepts such as influential and important are often read synonymously to centrality. There are a number of centrality measures for static networks, each capturing a different facet of centrality. Many of these measures can be translated to temporal networks fairly straightforwardly. For example, betweenness—roughly how many shortest paths that pass through a vertex—can be adapted to temporal networks by changing shortest paths to shortest time-respecting paths. It gets a bit more complicated with measures based on distances [33]. Closeness is defined as the reciprocal average distance from a node to the other nodes of a static graph. A simple way of dealing with missing time-respecting paths is to measure the average reciprocal distance (rather than the reciprocal average distance) and let a missing time-respecting path make a zero contribution to the average [12], [45]. Furthermore, as discussed in

[46], measures such as Katz centrality and PageRank can be generalized to temporal networks.

Another approach to identifying important spreaders is the vaccination problem. Assume you could vaccinate a fraction of a population for a future disease outbreak. Then, how can this fraction be chosen? This differs from the centrality approaches of finding important spreaders as it also takes the chance of getting the disease (or information, if we think of social media) into account. A celebrated approach in static networks is the neighbor vaccination method: pick persons at random, ask them to name an acquaintance (or someone they met when the disease in question could have spread), and then vaccinate the acquaintance. The probability the friend has a degree k is proportional to k; so high-degree individuals have a higher chance of being vaccinated. Lee *et al.* [47] proposed a version of this where one asks the randomly chosen person to name the most recent, or most frequent (some time in the past), acquaintance. Starnini *et al.* [48] proposed further improved methods. Such methods can exploit temporal heterogeneities and improve the efficiency of neighborhood vaccination.

F. Network Motifs

The idea behind network motifs comes from electronics. Small overrepresented subgraphs, primarily in directed networks, could be interpreted as building blocks in the network such as transistors in electronic circuitry [49]. In temporal networks, motifs have a bit different role. Rather, temporal network motifs are often thought of as common sequences of contacts among a small group of people. For example, Kovanen *et al.* [13] focus on contact sequences between vertices that are maximally separated by a time δt. Specifically, two contacts are said to be δt-adjacent if they share a vertex and are separated in time by δt or less. Pairs of contacts are then defined as δt-connected if there is a sequence of δt-adjacent events joining them. Then, Kovanen *et al.* proceed counting δt-connected subgraphs and comparing their frequency to those in randomized null models. They find an overrepresentation of subgraphs that seem to be causally connected (such as A contacts B who contacts C and D, as opposed to a noncausal sequence where B contacts C and D, and then A only contacts B). This work was motivated by cell phone call data, but its methodology could be straightforwardly applied to social media data. Also, not quite social media data, Jurgens and Lu [50] studied the evolution of Wikipedia by counting similar motifs. They associated common motifs with edit episodes such as "content reversion and antivandalism" or "collaborative editing."

G. Simplifying Temporal Networks

A final class of methods is how to simplify temporal network data. As mentioned above, it is hard to visualize temporal networks to give a feeling of the structure of even mid-sized data sets. For this reason, and also the more fundamental purpose to understand the important structures for spreading processes, one would need to simplify temporal network data. One approach is to project a temporal network to a static network. The straightforward way, to include an edge between all pairs of individuals that has at least one contact during the sampling period, is not always a good idea. If one is interested in a spreading process, that kind of projection could include too many irrelevant edges [51]. Holme [52] proposed either a carefully selected time window (and made a network of aggregated contacts within that window), or an exponential threshold representation where each contact contributes to an edge's weight with a term that is decaying exponentially from the start of the spreading process. This representation was shown to perform well to make a static network where the static network predictors of node importance for disease spreading match their actual importance in simulations directly on the temporal network.

Another type of projection to static graphs that could be useful, at least for very sparsely connected contact structures, is reachability graphs, or "path graphs" [25], or an "associated influence digraph" [53]. In these graphs, a directed edge from A to B means that (in the beginning of the data) there is a time-respecting path from A to B. This type of graphs tends to be extremely connected to real-world networks and thus not well adapted to complex network methods. See Fig. 4(a) for an illustration.

Instead of simplifying temporal networks to static networks one could project them to simpler forms of temporal networks. Holme and Liljeros [31] discuss "pictures" or components in models of temporal networks. In particular, two pictures are contrasted: a link turnover picture where one thinks of the first and last contacts of an edge as its beginning and end and ignores the timing of the other contacts; and an ongoing link picture where one thinks of the links as continuously active and the contacts as drawn from a probability distribution (perhaps reflecting a bursty interevent time statistics). Holme and Liljeros [31] argue that, with respect to disease spreading and empirical data sets, the link turnover picture is a better way of simplifying temporal networks. Fig. 4(b) and (c) illustrates the link turnover and ongoing link pictures, respectively. Ultimately, what determines which picture is most relevant depends on the relative time scales of the spreading processes and sampling time. If the sampling time is much shorter than spreading processes, then the ongoing link picture will be more relevant, and *vice versa*.

III. DISCUSSION AND CONCLUSION

Social media generate huge amount of metadata that could be used to understand social information flow, identify important spreaders, etc. Many kinds of such metadata could be represented as temporal networks, networks that record when contacts happen, in addition to who has been

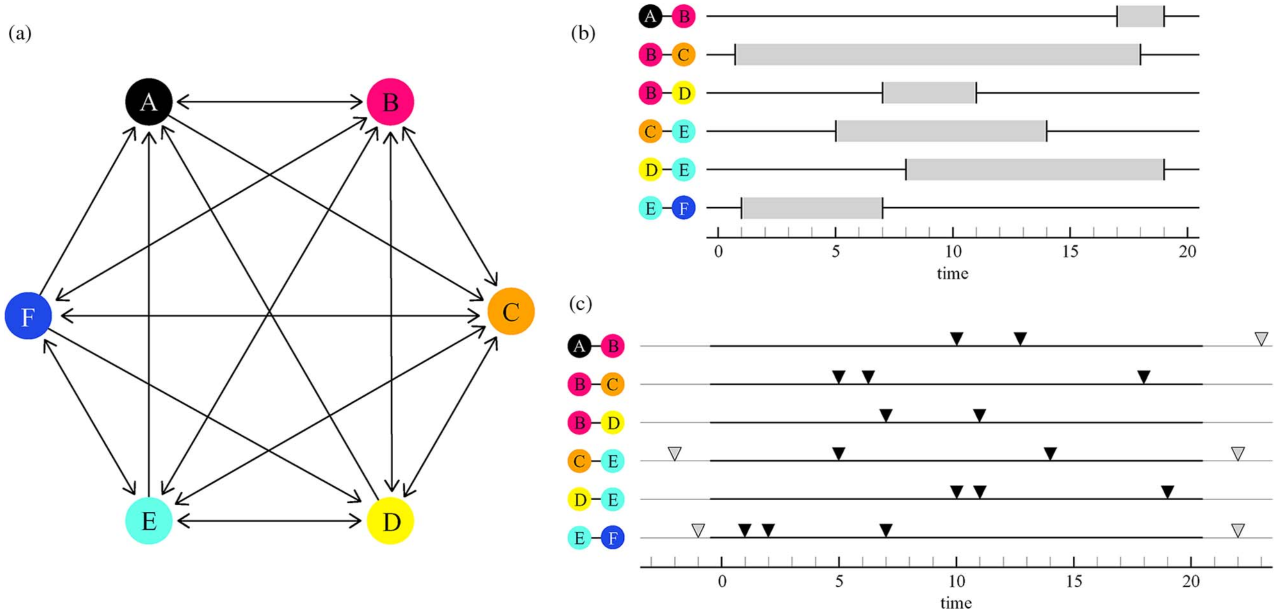


Fig. 4. ***Illustration of ways to simplify temporal network data. Panel A shows the reachability graph of the contact sequence displayed in Fig. 1. Panel B illustrates the link turnover picture as applied to the Fig. 1 network. In this picture, one thinks of the first observed contact as the beginning of the edge and the last contact as the end of the edge. Panel C illustrates the ongoing link picture where one assumes that the temporal network can be well described as an underlying static network with contacts happening over the edges following some interevent time distribution.***

in contact with whom. The field of temporal networks is still young and under rapid development. Probably, there are methods more specifically suited for social media data waiting to be developed. Nevertheless, there are a number of, primarily computational, techniques to analyze social media data sets that we discussed in this paper. These methods serve to identify important spreaders, characterize the behavior of the social media users, and map out temporal and topological structures. Recently, there has been much advancement in understanding of how such structures affect spreading phenomena [54]–[56]. It would be interesting to validate such theories using social media data. This would be happening in a different way than usual—inferring predictors of spreading cascades from data [42], [57], [58]. Both fields of temporal networks and theory of social media need each other; the former needs real data to test and develop theories, and the latter needs a theoretical framework to handle large time-resolved data.

Word-of-mouth marketing and the detection and surveillance of infectious disease are two areas where temporal network methods have a great potential of improving existing methods [59]. These are areas where researchers already have applied static network methods to a great extent. At the same time, the underlying interaction has a strong temporal component, meaning that temporal network techniques could potentially be very fruitful. For example, Christakis and Fowler [60] propose to use neighbors of random individuals as sentinels for detection of epidemic outbreaks. Neighbors of average nodes are more central than average [61], which could be exploited. Christakis and Fowler do not discuss the use of social media data as such, but social networks in general. However, finding infected individuals from social media data is not difficult *per se* [62]–[65], so their method seems possible to straightforwardly extend to this type of data, and possible to improve by methods like the ones discussed in this paper (Sections II-E and II-G). Social media data could also be used to monitor the sentiments toward public health programs [66], [67]. Temporal network methods could be used to analyze or model the opinion spreading behind changes in this type of sentiments.

Another interesting application of temporal networks to social media data could be network interventions in the process of "using social network data to accelerate behavior change or improve organizational performance" [68]. A typical application would be to identify individuals or groups whose change of behavior can trigger a cascade of behavioral change. Another typical task is to find undesired grouping, splits, or hierarchical dependencies in the social network of an organization, and then find a way to improve the situation by reorganizing. These social networks that network interventions rely on are not static and in the dynamic aspects of their nature lies much information that could be exploited by methods described in this paper.

In this paper, we have tried to argue that temporal network techniques are well suited for social media data. Compared to many other areas within the umbrella term of network science (including traditional social network analysis, focusing on [69] and [70]), social media data are accurately time tagged, and thus readily analyzable by temporal network methods. We have mentioned how temporal structures can change dynamics of information spreading events. Clearly, without using temporal information, prediction and modeling would be less precise, just as the static network structure can add precision compared to well-mixed models. Not only that, one could miss the most important ways to estimate the spreading speed, decide who is the most important (or influential) individual, find efficient ways to mitigate or enhance spreading, etc. Temporal network approaches to social media are, we believe, an understudied area, so we expect much more research in this direction in the near future. ■

Acknowledgment

The author would like to thank J. Saramäki and N. Masuda for discussions.

REFERENCES

[1] P. Holme and J. Saramäki, "Temporal networks," *Phys. Rep.*, vol. 529, pp. 97–125, 2012.

[2] J.-P. Eckmann, E. Moses, and D. Sergi, "Entropy of dialogues creates coherent structures in e-mail traffic," *Proc. Nat. Acad. Sci. USA*, vol. 101, pp. 14333–14337, 2004.

[3] H. Ebel, H.-I. Mielsch, and S. Bornholdt, "Scale-free topology of e-mail networks," *Phys. Rev. E*, vol. 66, 2002, 035103.

[4] B. Viswanath, A. Mislove, M. Cha, and F. P. Gummadi, "On the evolution of user interaction in Facebook," in *Proc. 2nd ACM Workshop Online Social Netw.*, 2009, pp. 37–42.

[5] S. A. Golder, D. M. Wilkinson, and B. A. Huberman, "Rhythms of social interaction: Messaging within a massive online network," in *Communities and Technologies 2007*, C. Steinfield, B. T. Pentland, M. Ackerman, and N. Contractor, Eds. London, U.K.: Springer-Verlag, 2007, pp. 41–66.

[6] E. Gilbert and K. Karahalios, "Predicting tie strength with social media," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2009, pp. 211–220.

[7] P. Holme, C. R. Edling, and F. Liljeros, "Structure and time-evolution of an internet dating community," *Social Netw.*, vol. 26, pp. 155–174, 2004.

[8] F. Karimi, V. C. Ramenzoni, and P. Holme, "Structural differences between open and direct communication in an online community," *Physica A*, vol. 414, pp. 263–273, 2014.

[9] J. Leskovec, D. Huttenlocher, and J. Kleinberg, "Signed networks in social media," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2010, pp. 1361–1370.

[10] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?" in *Proc. 19th Int. World Wide Web Conf.*, Raleigh, NC, USA, Apr. 26–30, 2010, pp. 591–600.

[11] A. Noulas, S. Scellato, C. Mascolo, and M. Pontil, "An empirical study of geographic user activity patterns in Foursquare," in *Proc. 5th Int. AAAI Conf. Weblogs Social Media*, Barcelona, Spain, 2011, pp. 570–573.

[12] R. K. Pan and J. Saramäki, "Path lengths, correlations, centrality in temporal networks," *Phys. Rev. E*, vol. 84, 2011, 016105.

[13] L. Kovanen, M. Karsai, K. Kaski, J. Kertész, and J. Saramäki, "Temporal motifs in time-dependent networks," *J. Stat. Mech.*, 2011, DOI: doi:10.1088/1742-5468/2011/11/P11005

[14] M. Karsai *et al.*, "Small but slow world: How network topology and burstiness slow down spreading," *Phys. Rev. E*, vol. 83, 2011, 025102.

[15] J. Leskovec, L. Backstrom, and J. Kleinberg, "Meme-tracking and the dynamics of the news cycle," in *Proc. 15th ACM SIGKDD Int. Conf. Knowl. Disc. Data Mining*, 2009, pp. 497–506.

[16] D. J. Bartholomew, *Stochastic Models for Social Processes*. New York, NY, USA: Wiley, 1967.

[17] R. M. Anderson and R. M. May, *Infectious Diseases of Humans: Dynamics and Control.* Oxford, U.K.: Oxford Univ. Press, 1991.

[18] P. Clifford and A. W. Sudbury, "A model for spatial conflict," *Biometrika*, vol. 60, pp. 581–588, 1973.

[19] T. M. Liggett, *Interacting Particle Systems.* New York, NY, USA: Springer-Verlag, 1985.

[20] M. Granovetter, "Threshold models of collective behavior," *Amer. J. Sociol.*, vol. 83, pp. 1420–1443, 1978.

[21] F. Karimi and P. Holme, "Threshold model of cascades in empirical temporal networks," *Physica A*, vol. 392, pp. 3476–3483, 2013.

[22] T. Takaguchi, N. Masuda, and P. Holme, "Bursty communication patterns facilitate spreading in a threshold-based epidemic dynamics," *PLoS ONE*, vol. 8, 2013, e68629.

[23] D. Kempe, J. Kleinberg, and A. Kumar, "Connectivity and inference problems for temporal networks," in *Proc. 32nd Annu. ACM Symp. Theory Comput.*, 2000, pp. 504–513.

[24] P. Holme, "Network reachability of real-world contact sequences," *Phys. Rev. E*, vol. 71, 2005, 046119.

[25] J. Moody, "The importance of relationship timing for diffusion," *Social Forces*, vol. 81, pp. 25–56, 2002.

[26] K. L. Cooke and E. Halsey, "The shortest route through a network with time-dependent internodal transit times," *J. Math. Anal. Appl.*, vol. 14, pp. 493–498, 1966.

[27] L. Lamport, "Time, clocks, the ordering of events in a distributed system," *Commun. ACM*, vol. 21, pp. 558–565, 1978.

[28] F. Mattern, "Virtual time and global states of distributed systems," in *Proc. Workshop Parallel Distrib. Algorithms*, 1989, pp. 215–226.

[29] C. Lee, B. Nick, U. Brandes, and P. Cunningham, "Link prediction with social vector clocks," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Disc. Data Mining*, 2013, pp. 784–792.

[30] G. Miritello, R. Lara, M. Cebrian, and E. Moro, "Limited communication capacity unveils strategies for human interaction," *Sci. Rep.*, vol. 3, 2013, 1950.

[31] P. Holme and F. Liljeros, "Birth and death of links control disease spreading in empirical contact networks," *Sci. Rep.*, vol. 4, 2014, 4999.

[32] G. Kossinets, J. Kleinberg, and D. J. Watts, "The structure of information pathways in a social communication network," in *Proc. 14th ACM SIGKKD Int. Conf. Knowl. Disc. Data Mining*, 2008, pp. 435–443.

[33] J. Tang, M. Musolesi, C. Mascolo, and V. Latora, "Temporal distance metrics for social network analysis," in *Proc. 2nd ACM SIGCOMM Workshop Online Social Netw.*, 2009, pp. 31–36.

[34] V. Kostakos, "Temporal graphs," *Physica A*, vol. 388, pp. 1007–1023, 2009.

[35] L. E. C. Rocha, F. Liljeros, and P. Holme, "Simulated epidemics in an empirical spatiotemporal network of 50, 185 sexual contacts," *PLoS Comput. Biol.*, vol. 7, 2011, e1001109.

[36] W. E. Leland and D. V. Wilson, "High time-resolution measurement and analysis of LAN traffic: Implications for LAN interconnection," in *Proc. InfoCom*, 1991, pp. 1360–1366.

[37] P. Holme, "N2etwork dynamics of ongoing social relationships," *Europhys. Lett.*, vol. 64, pp. 427–433, 2003.

[38] H.-H. Jo, R. K. Pan, and K. Kaski, "Emergence of bursts and communities in evolving weighted networks," *PLoS ONE*, vol. 6, 2011, e22687.

[39] T. Zhou, Z.-D. Zhao, Z. Yang, and C. Zhou, "Relative clock verifies endogenous bursts of human dynamics," *Europhys. Lett.*, vol. 97, 2012, 18006.

[40] P. Bajardi, A. Barrat, F. Natale, L. Savini, and V. Colizza, "Dynamical patterns of cattle trade movements," *PLoS ONE*, vol. 6, 2011, e19869.

[41] P. Holme, C. E. Edling, and F. Liljeros, "Structure and time-evolution of an Internet dating community," *Social Netw.*, vol. 26, pp. 155–174, 2004.

[42] S. Aral and D. Walker, "Identifying influential and susceptible members of social networks," *Science*, vol. 337, pp. 337–341, 2012.

[43] J. Leskovec, L. A. Adamic, and B. A. Huberman, "The dynamics of viral marketing," *ACM Trans. Web*, vol. 1, 2007, article 5.

[44] P. Basaras, D. Katsaros, and L. Tassiulas, "Detecting influential spreaders in complex, dynamic networks," *Computer*, vol. 46, pp. 24–29, 2013.

[45] N. Santoro, W. Quattrociocchi, P. Flocchini, A. Casteigts, and F. Amblard, "Time-varying graphs and social network analysis: Temporal indicators and metric," in *Proc. 3rd AISB Social Netw. Multiagent Syst. Symp.*, 2011, pp. 32–38.

[46] P. Grindrod, M. C. Parsons, D. J. Higham, and E. Estrada, "Communicability across evolving networks," *Phys. Rev. E*, vol. 81, 2011, 046120.

[47] S. Lee, L. E. C. Rocha, F. Liljeros, and P. Holme P, "Exploiting temporal network structures of human interaction to effectively immunize populations," *PLoS ONE*, vol. 7, 2012, e36439.

[48] M. Starnini, A. Machens, C. Cattuto, A. Barrat, and R. Pastor-Satorras, "Immunization strategies for epidemic processes in time-varying contact networks," *J. Theor. Biol.*, pp. 89–100, 2013, article 337.

[49] U. Alon, "Network motifs: Theory and experimental approaches," *Nature Rev. Genetics*, vol. 8, pp. 450–461, 2007.

[50] D. Jurgens and T.-S. Lu, "Temporal motifs reveal the dynamics of editor interactions in Wikipedia," in *Proc. 6th Int. AAAI Conf. Weblogs Social Media*, 2012, pp. 162–169.

[51] G. Krings, M. Karsai, S. Bernhardsson, V. D. Blondel, and J. Saramäki, "Effects of time window size and placement on the structure of an aggregated communication network," *EPJ Data Sci.*, vol. 1, 2012, article 4.

[52] P. Holme, "Epidemiologically optimal static networks from temporal network data," *PLoS Comput. Biol.*, vol. 9, 2013, e1003142.

[53] E. Cheng, J. W. Grossman, and M. J. Lipman, "Time-stamped graphs and their associated influence digraphs," *Disc. Appl. Math.*, vol. 128, pp. 317–335, 2013.

[54] N. Masuda, K. Klemm, and V. M. Eguíluz, "Temporal networks: Slowing down diffusion by long lasting interactions," *Phys. Rev. Lett.*, vol. 111, 2013, 188701.

[55] R. Lambiotte, L. Tabourier, and J.-C. Delvenne, "Burstiness and spreading on temporal networks," *Eur. Phys. J. B*, vol. 86, 2013, article 320.

[56] H.-H. Jo, J. I. Perotti, K. Kaski, and J. Kertész, "Analytically solvable model of spreading dynamics with non-Poissonian processes," *Phys. Rev. X*, vol. 4, 2014, 011041.

[57] L. Weng, F. Menczer, and Y.-Y. Ahn, "Virality prediction and community structure in social networks," *Sci. Rep.*, vol. 3, 2013, article 2522.

[58] J. Cheng, L. A. Adamic, P. A. Dow, J. Kleinberg, and J. Leskovec, "Can cascades be predicted?" in *Proc. World Wide Web Conf.*, Seoul, Korea, 2014, DOI: 10.1371/journal.pcbi.1002199

[59] P. Holme and M. Huss, "Understanding and exploiting information spreading and integrating information technologies," *J. Comput. Sci. Technol.*, vol. 26, pp. 829–836, 2011.

[60] N. A. Christakis and J. H. Fowler, "Social network sensors for early detection of contagious outbreaks," *PLoS ONE*, vol. 5, 2010, e12948.

[61] M. E. J. Newman, "Ego-centered networks and the ripple effect," *Social Netw.*, vol. 25, pp. 83–95, 2003.

[62] V. Lampos, T. de Bie, and N. Cristianini, "Flu detector: Tracking epidemics on Twitter," in *Proc. Eur. Conf. Mach. Learn. Principles Practice Knowl. Disc. Databases*, 2010, pp. 599–602.

[63] A. Signorini, A. M. Segre, and P. M. Polgreen, "The use of Twitter to track levels of disease activity and public concern in the U.S. during the influenza a H1N1 pandemic," *PLoS ONE*, vol. 6, 2011, e19467.

[64] T. Bodnar and M. Salathé, "Validating models for disease detection using twitter," in *Proc. World Wide Web Conf.*, 2013, pp. 699–702.

[65] E. Coeira, "Social networks, social media, social diseases," *Brit. Med. J.*, vol. 346, 2013, f3007.

[66] M. Salathé and S. Khandelwal, "Assessing vaccination sentiments with online social media: Implications for infectious disease dynamics and control," *PLoS Comput. Biol.*, vol. 7, 2011, e1002199.

[67] M. Salathé, D. Q. Vu, S. Khandelwal, and D. R. Hunter, "The dynamics of health behavior sentiments on a large online social network," *EPJ Data Sci.*, vol. 2, 2013, article 4.

[68] T. W. Valente, "Network interventions," *Science*, vol. 337, pp. 49–53, 2012.

[69] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 1994.

[70] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, "Network analysis in the social sciences," *Science*, vol. 323, pp. 892–895, 2009.

ABOUT THE AUTHOR

Petter Holme received the M.Sc. degree in engineering physics from Uppsala University, Uppsala, Sweden, the M.A. degree in Chinese from Stockholm University, Stockholm, Sweden, and the Ph.D. degree in theoretical physics from Umeå University, Umeå, Sweden.

He was a Postdoctoral Fellow at the Department of Physics, University of Michigan, Ann Arbor, MI, USA and the Department of Computer Science, University of New Mexico, Albuquerque, NM, USA; an Assistant Professor at the Royal Institute of Technology, Stockholm, Sweden; and an Associate Professor at Umeå University. Currently, he is an Associate Professor at Sungkyunkwan University, Suwon, Korea. He has around 100 publications spanning various topics, mostly using networks as a modeling framework. His current main research is how to integrate temporal and spatial information to improve models of spreading phenomena on networks.

INVITED PAPER

Tracking the Digital Footprints of Personality

This paper reviews literature showing how pervasive records of digital footprints can be used to infer personality.

By Renaud Lambiotte and Michal Kosinski

ABSTRACT | A growing portion of offline and online human activities leave digital footprints in electronic databases. Resulting big social data offers unprecedented insights into population-wide patterns and detailed characteristics of the individuals. The goal of this paper is to review the literature showing how pervasive records of digital footprints, such as Facebook profile, or mobile device logs, can be used to infer personality, a major psychological framework describing differences in individual behavior. We briefly introduce personality and present a range of works focusing on predicting it from digital footprints and conclude with a discussion of the implications of these results in terms of privacy, data ownership, and opportunities for future research in computational social science.

KEYWORDS | Big data; personality; psychology; social networks

I. INTRODUCTION

In recent years, a growing portion of human activities such as social interactions and entertainment have become mediated by digital services and devices. The records of those activities, or "big social data," are changing the paradigm in the social sciences, as it undergoes a transition from small-scale studies, typically employing questionnaires or lab-based observations and experiments, to large-scale studies, in which researchers observe the behavior of thousands or millions of individuals and search for statistical regularities and underlying principles [1]–[6]. These works provide empirical observations at an unprecedented scale offering the potential to radically improve our understanding of the individuals and social systems.

One of the major insights offered by big social data research relates to the predictability of individuals' psychological traits from their digital footprint [3]. Ability to automatically assess psychological profiles opens the way for improved products and services as personalized search engines, recommender systems [7], and targeted online marketing [8]. On the other hand, however, it creates significant challenges in the areas of privacy [9], [10]. The main goal of this paper is to provide a review of the works investigating the potential of the big social data to predict a five-factor model of personality—the major set of psychological traits—supporting further studies of the relationship between personality and digital footprint and its implications for privacy and new products and services.

II. PERSONALITY

The most widespread and generally accepted model of personality is the five-factor model of personality (FFM; [11]). FFM was shown to subsume most known personality traits, and it is claimed to represent the basic structure underlying the variations in human behavior and preferences, providing a nomenclature and a conceptual framework that unifies much of the research findings in the psychology of individual differences. FFM includes the following traits.

1) Openness is related to imagination, creativity, curiosity, tolerance, political liberalism, and appreciation for culture. People scoring high on openness like change, appreciate new and unusual ideas, and have a good sense of aesthetics.

Manuscript received January 29, 2014; revised July 24, 2014; accepted September 9, 2014. Date of publication October 29, 2014; date of current version November 18, 2014. The work of R. Lambiotte was supported by the F.R.S.-Fonds de la Recherche Scientifique (FNRS), the European Union (EU) project Optimizr, and COST Action TD1210 KnowEscape. The work of M. Kosinski was supported by the Psychometrics Centre at the University of Cambridge, Boeing Corporation, Microsoft Research, the National Science Foundation (NSF), the Defense Advanced Research Projects Agency (DARPA), and Center for the Study of Language and Information at Stanford University (CLSI). This paper presents results of the Belgian Network Dynamical Systems, Control, and Optimization (DYSCO), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office.

R. Lambiotte is with the Namur Center for Complex Systems (naXys), University of Namur, Namur 5000, Belgium (e-mail: renaud.lambiotte@unamur.be).
M. Kosinski is with InfoLab, Stanford University, Stanford, CA 94305 USA, and also with the Psychometrics Centre, University of Cambridge, Cambridge CB2 1TN, U.K.

Digital Object Identifier: 10.1109/JPROC.2014.2359054

2) Conscientiousness measures the preference for an organized approach to life in contrast to a spontaneous one. Conscientious people are more likely to be well organized, reliable, and consistent. They enjoy planning, seek achievements, and pursue long-term goals. Nonconscientious individuals are generally more easygoing, spontaneous, and creative. They tend to be more tolerant and less bound by rules and plans.
3) Extroversion measures a tendency to seek stimulation in the external world, the company of others, and to express positive emotions. Extroverts tend to be more outgoing, friendly, and socially active. They are usually energetic and talkative; they do not mind being at the center of attention and make new friends more easily. Introverts are more likely to be solitary or reserved and seek environments characterized by lower levels of external stimulation.
4) Agreeableness relates to a focus on maintaining positive social relations, being friendly, compassionate, and cooperative. Agreeable people tend to trust others and adapt to their needs. Disagreeable people are more focused on themselves, less likely to compromise, and may be less gullible. They also tend to be less bound by social expectations and conventions and are more assertive.
5) Emotional stability (opposite referred to as neuroticism) measures the tendency to experience mood swings and emotions, such as guilt, anger, anxiety, and depression. Emotionally unstable (neurotic) people are more likely to experience stress and nervousness, whereas emotionally stable people (low neuroticism) tend to be calmer and self-confident.

Research has shown that personality is correlated with many aspects of life, including job success [12], attractiveness [13], drug use [14], marital satisfaction [15], infidelity [16], and happiness [17]. The main limitations of classical personality studies are, however, the size of the samples, often too poor for statistical validation, and their strong bias toward white, educated, industrialized, rich, and democratic (WEIRD) people [18].

III. FROM OFFLINE TO ONLINE...

The increasingly prevalent access to digital media enables large-scale online projects aimed at collecting personality profiles and exploring their relations with digital footprints. Personality has been investigated through different types of online media, for instance, by focusing on website browsing logs [2], [19], contents of personal websites [20], music collections [21], or properties of Twitter profiles [22], [23].

The most complete online social environment is arguably Facebook, due to its popularity and rich social and semantic data stored on its users' profiles that can be conveniently recorded. It is important to note that Facebook profiles are increasingly becoming a channel through which to form impressions about others, for example, before dating [24] or before a job interview [25]. Moreover, research tends to show that a Facebook profile reflects the actual personality of an individual rather than an idealized role [26], and that personality can be successfully judged by the others based on Facebook profiles [27], [28]. These results suggest that personality is manifested not only in the offline, but also online behavior, and thus digital footprints can be used to predict it.

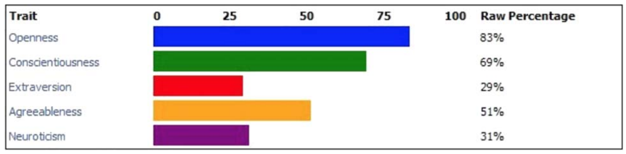


Fig. 1. ***Snapshot of a personality profile generated by the myPersonality Facebook App, representing an individual that is liberal and open minded (high openness), well-organized (high conscientiousness), contemplative and happy with own company (low extroversion), of average competitiveness (average agreeableness), and laid back and relaxed (low neuroticism).***

The most popular data set used to study the relationship between personality and digital footprint comes from the myPersonality project. myPersonality was a Facebook application set up by David Stillwell in 2007 that offered participants access to 25 psychological tests and attracted over six million users. myPersonality users received immediate feedback (see Fig. 1) on their results and could donate their Facebook profile information to research resulting in a database that, after anonymization, is being shared with the academic community at mypersonality.org, allowing for the study of hitherto unanswered questions in a wide range of topics, such as geographical variations in personality ([29]; see Fig. 2), social networks [2], [22], [30], [31], privacy [32], language [6] (see Fig. 3), predicting individual traits [33], [3], computer science [34], happiness [35], music [36], and delayed discounting [37].

IV. SOCIAL NETWORK STRUCTURE

Social network structure is one of the major types of digital footprint left by the users, and a growing number of studies shows that it is predictive of often intimate personal traits. For instance, it is known that the location within a Facebook friendship network is predictive of sexual orientation [38]. Similarly, it is possible to accurately detect users' romantic partner by observing overlap in social circles [39].

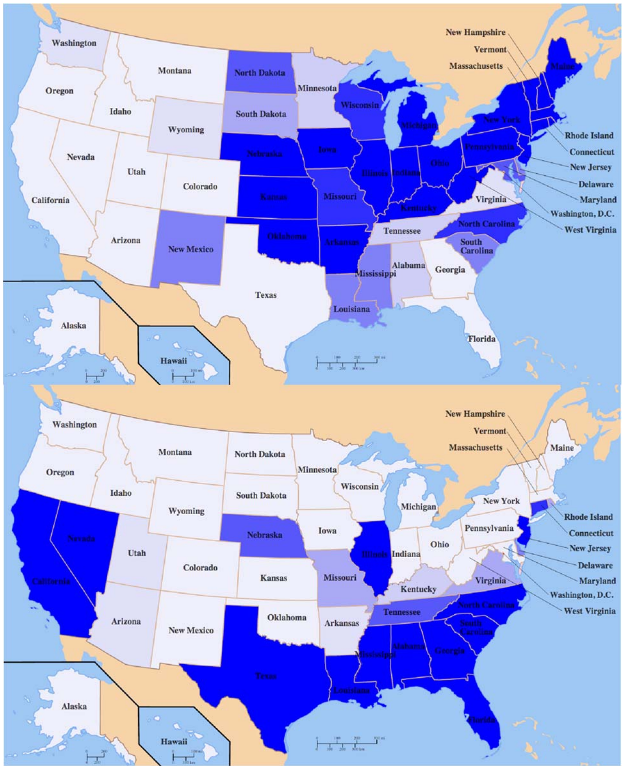


Fig. 2. *Personality maps of U.S. states for neuroticism (upper) and extroversion (lower). Dark (light) blue indicates values higher (lower) than average. Figure based on myPersonality data.*

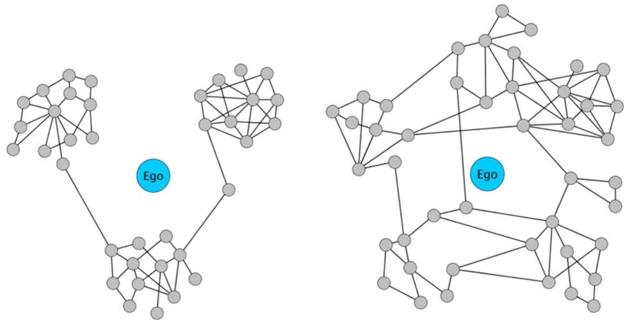


Fig. 4. *Typical egocentric networks of introverts (left) and extroverts (right). Introverts tend to belong to fewer but larger and denser communities, while extroverts tend to act as bridges between more frequent, smaller, and overlapping communities. Connections between Ego and his friends have not been depicted for the sake of clarity.*

Personality is expected to affect people's social network surroundings as it affects the types and number of social ties formed by people. There are a number of studies exploring this relationship. Neuroticism is usually associated with negative social interactions, while extroversion positively correlates with the size of the network and greater social status [40], [41]. Results related to the remaining traits tend to be inconsistent, perhaps due to small sample sizes. More recently, Quercia *et al.* [31] used myPersonality data set to study the relation between sociometric popularity and personality traits, at a scale several orders of magnitudes larger than in the previous studies. They have shown that the strongest predictor for the number of friends is extroversion, while other personality traits do not play a significant role. On average, extreme extroverts tend to have twice as many friends as extreme introverts. A subsequent work [42] went one step further and, for the first time, quantitatively explained the way in which egocentric network topology is shaped by personality. It confirmed that extroversion plays a major role by showing that introverts are part of fewer but larger communities, whereas extroverts tend to act as bridges between more frequent but smaller communities (see Fig. 4).

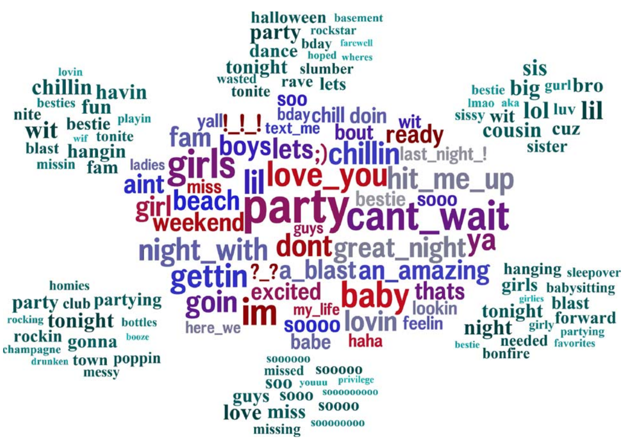


Fig. 3. *Words, phrases, and topics most distinguishing extroversion from introversion. Source: [6].*

V. FACEBOOK LIKES

The Facebook profile of a user is not purely demographic, as it also contains robust records of digital footprints. In particular, Facebook likes exemplify a typical variety of digital footprint—a connection between the user and a content that is similar to other pervasive records such as playlists (see Fig. 5), website browsing logs, purchase records, or web search queries. A recent paper [3] based on the myPersonality database and using relatively straightforward methods (singular value decomposition and linear regression) showed that Facebook likes are highly predictive of personality and number of other psychodemographic traits, such as age, gender, intelligence, political and religious views, and sexual orientation (see Fig. 6). The paper provided examples of likes most strongly associated with given personality traits. For example, users who liked "Hello Kitty" brand tended to

Fig. 5. *Dendrogram illustrating the structure of music tastes and its relationship to the personality trait of openness among myPersonality users. The structure was produced using hierarchical clustering of the most popular Facebook likes from musician/band category. The color scale represents the average openness of its subscribers, ranging from conservative (cyan) to liberal (magenta). The height of the nodes is proportional to the dissimilarity between individual likes or clusters at both ends. The shorter is the path between two musicians or bands, the larger overlap in audience. Source: [43].*

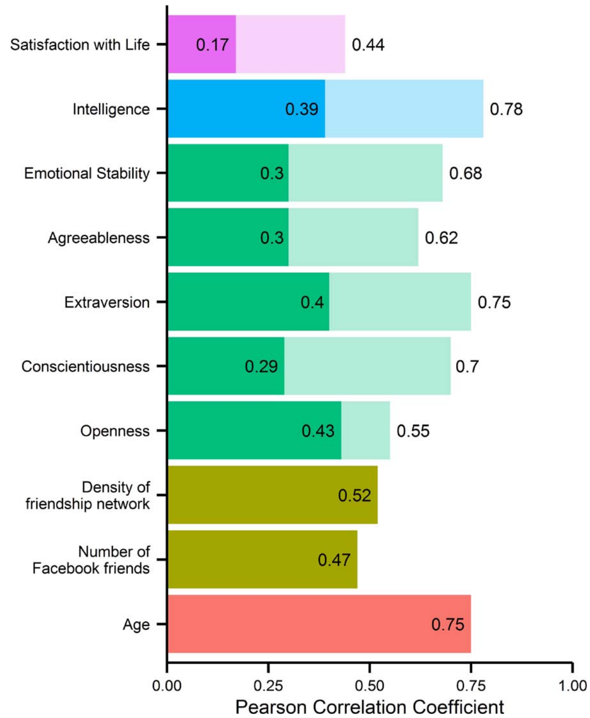


Fig. 6. *Prediction accuracy of regression for numeric attributes and traits expressed by the Pearson correlation coefficient between predicted and actual attribute values; all correlations are significant at the $p < 0.001$ level. The red outline bars indicate the questionnaire's baseline accuracy, expressed in terms of test-retest reliability. Source: [3].*

have high openness, low conscientiousness, and low agreeableness.

VI. SEMANTIC ANALYSIS

Similar predictions can be based on the textual analysis of people's posts and other samples of text. There is a long tradition in using text to infer personality [44], [45], [46], however, never at the scale presented in [6]. This study applied differential language analysis to uncover features distinguishing demographic and psychological attributes to 700 million words, phrases, and topic instances collected by myPersonality from Facebook status updates of 75 000 participants. It showed a striking variations of language driven by personality, gender, and age. This work has not only confirmed existing observations (such as neurotic people's tendency to use the word "depressed"), but also posed new hypotheses (such as a relationship between physical activity and low neuroticism).

VII. . . . AND BACK FROM ONLINE TO OFFLINE

The proliferation of mobile-devices loaded with sensors means that offline human activities are also increasingly leaving digital footprint [47], [48]. For instance, physical states such as running or walking can be inferred from accelerometer data; colocation with other devices can be detected using Bluetooth; geolocation can be established using WiFi, Global Positioning System (GPS), or Global System for Mobile (GSM) triangulation; and social interactions can be measured by records of text messages and phone calls. These data can be recorded by dedicated apps, such as EmotionSense [49], which measures emotional states based on the speech patterns and matches it with physical activity, geolocation, and colocation with other users. In the last few years, call data records (CDRs) have been used to study the organization of social networks and human mobility [50], [51], [52].

Similarly to digital footprints left in the online environment, offline activities recorded with mobile devices' sensors reflect users' personality. A recent study combined CDRs with personality profiles of mobile device users and identified a number of mobility and social factors correlated with personality [53]. For instance, mobility indicators, such as distance traveled, significantly correlate with neuroticism, while social life indicators, such as the size of the social network, correlated with extroversion, in agreement with the previous results based on online digital footprints.

VIII. CONCLUSION

The main purpose of this paper was to review the evidence of the relationship between digital footprint and personality. We have shown that a wide range of pervasive and often publicly available digital footprints such as Facebook profiles or data from mobile devices can be used to infer personality. As our life is increasingly interwoven with digital services and devices, it is becoming critical to understand the consequences of the apparent ability to automatically and rapidly assess people's psychological traits.

Works cited in this paper indicate that the accuracy of the personality predictions is moderate, with typical correlation between the prediction and personality in the range of $r = 0.2$ and $r = 0.4$. It has to be noted, however, that the ground truth (i.e., personality scores) is also merely an approximation of the underlying latent traits. For example, the accuracy of the personality scales used in [3] expressed as a correlation between scores achieved by the same person in two points of time (test-retest reliability) ranged between $r = 0.55$ and $r = 0.75$. It is reasonable to expect that with, an increasing amount of data available and improved methods, assessment accuracy will improve.

Predicting users' personality can be used to improve numerous products and services. Digital systems and devices (such as online stores or cars) could be designed to adjust their behavior to best fit their users' inferred profiles [54]. For example, a car could adjust the parameters of the engine and the music to the personality and current mood of the driver. Also, the relevance of marketing and product recommendations could be improved by adding psychological dimensions to current user models. For example, online insurance advertisements might emphasize security when facing emotionally unstable (neurotic) users but stress potential threats when dealing with emotionally stable ones. Moreover, digital footprint may provide a convenient and reliable way to measure psychological traits at a low cost. Such automated assessment could prove to be more accurate and less prone to cheating and misrepresentation than traditional questionnaires.

Furthermore, it is likely that new insights into individual differences in human behavior offered by big social data will fuel the emergence of new, more accurate, robust models describing individuals and societies [5]. The translation of big social data into models and policies calls for a new wave of multidisciplinary collaborations between fields as diverse as psychology, social sciences, linguistics, computer science, and applied mathematics (perhaps under the banner of computational social psychology).

On the other hand, the results presented here may have considerable negative implications because it can easily be applied to large numbers of people without obtaining their individual consent and without them noticing. Commercial companies, governmental institutions, or even one's Facebook friends could use software to infer personality (and other attributes, such as intelligence or sexual orientation) that an individual may not have intended to share. There is a risk that the growing awareness of such digital exposure may decrease their trust in digital technologies, or even completely deter them from them. We hope that researchers, policy makers, and customers will find solutions to address those challenges and retain the balance between the promises and perils of the Digital Age. ■

REFERENCES

[1] N. Aharony, W. Pan, C. Ip, I. Khayal, and A. Pentland, "The social FMRI: Measuring, understanding, designing social mechanisms in the real world," in *Proc. 13th Int. Conf. Ubiquitous Comput.*, 2011, pp. 445–454.

[2] M. Kosinski, Y. Bachrach, P. Kohli, D. Stillwell, and T. Graepel, "Manifestations of user personality in website choice and behaviour on online social networks," *Mach. Learn.*, vol. 95, pp. 357–380, 2013.

[3] M. Kosinski, D. Stillwell, and T. Graepel, "Private traits and attributes are predictable from digital records of human behavior," *Proc. Nat. Acad. Sci.*, vol. 110, pp. 5802–5805, 2013.

[4] A. Kramer, J. Guillory, and J. Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *Proc. Nat. Acad. Sci.*, vol. 111, pp. 8788–8790, 2014.

[5] D. Lazer *et al.*, "Social science: Computational social science," *Science*, vol. 323, no. 5915, pp. 721–723, 2009.

[6] H. Schwartz *et al.*, "Personality, gender, age in the language of social media: The open-vocabulary approach," *PloS One*, vol. 8, no. 9, 2013, DOI: 10.1371/journal.pone.0073791.

[7] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, 2009.

[8] Y. Chen, D. Pavlovy, and J. F. Canny, "Large-scale behavioral targeting," in *Proc. Conf. Knowl. Disc. Data Mining*, 2009, pp. 209–217.

[9] D. Butler, "Data sharing threatens privacy," *Nature*, vol. 449, no. 7163, pp. 644–645, 2007.

[10] A. Narayanan and V. Shmatikov, "Robust de-anonymization of large sparse datasets," in *Proc. Symp. Security Privacy*, 2008, pp. 111–125.

[11] L. Goldberg, "The structure of phenotypic personality traits," *Amer. Psychol.*, vol. 48, no. 1, pp. 26–34, 1993.

[12] T. Judge, C. Higgins, C. Thoresen, and M. Barrick, "The Big Five personality traits, general mental ability, career success across the life span," *Personnel Psychol.*, vol. 52, no. 3, pp. 621–652, 1999.

[13] D. Byrne, W. Griffitt, and D. Stefaniak, "Attraction and similarity of personality characteristics," *J. Personality Social Psychol.*, vol. 5, no. 1, pp. 82–90, 1967.

[14] B. W. Roberts, O. S. Chernyshenko, S. Stark, and L. R. Goldberg, "The structure of conscientiousness: An empirical investigation based on seven major personality questionnaires," *Personnel Psychol.*, vol. 58, no. 1, pp. 103–139, 2005.

[15] E. Kelly and J. Conley, "Personality and compatibility: A prospective analysis of marital stability and marital satisfaction," *J. Personality Social Psychol.*, vol. 52, no. 1, pp. 27–40, 1987.

[16] T. Orzeck and E. Lung, "Big-Five personality differences of cheaters and non-cheaters," *Current Psychol.*, vol. 24, pp. 274–287, 2005.

[17] D. Ozer and V. Benet-Martinez, "Personality and the prediction of consequential outcomes," *Annu. Rev. Psychol.*, vol. 57, pp. 401–421, 2006.

[18] J. Henrich, S. Heine, and A. Norenzayan, "The WEIRDest people in the world," *Behav. Brain Sci.*, vol. 33, no. 2–3, pp. 61–83, 2010.

[19] J. Hu, H.-J. Zeng, H. Li, C. Niu, and Z. Chen, "Demographic prediction based on user's browsing behavior," in *Proc. Int. World-Wide Web Conf.*, 2007, pp. 151–160.

[20] B. Marcus, F. Machilek, and A. Schütz, "Personality in cyberspace: Personal web sites as media for personality expressions and impressions," *J. Personality Social Psychol.*, vol. 90, no. 6, pp. 1014–1031, 2006.

[21] P. Rentfrow and S. Gosling, "The do re mi's of everyday life: The structure and personality correlates of music preferences," *J. Personality Social Psychol.*, vol. 84, no. 6, pp. 1236–1256, 2003.

[22] D. Quercia, M. Kosinski, D. Stillwell, and J. Crowcroft, "Our Twitter profiles, our selves: Predicting personality with Twitter,"

in *Proc. Int. Conf. Privacy Security Risk Trust/IEEE Int. Conf. Social Comput.*, 2011, pp. 180–185.

[23] J. Golbeck, C. Robles, M. Edmondson, and K. Turner, "Predicting personality from Twitter," in *Proc. Int. Conf. Privacy Security Risk Trust/IEEE Int. Conf. Social Comput.*, 2011, pp. 149–156.

[24] S. Zhao, S. Grasmuck, and J. Martin, "Identity construction on Facebook: Digital empowerment in anchored relationships," *Comput. Human Behav.*, vol. 24, no. 5, pp. 1816–1836, 2008.

[25] A. Finder, "For some, online persona undermines a résumé," *New York Times*, Jun. 2006. [Online]. Available: http://www.nytimes.com/2006/06/11/us/11recruit.html?pagewanted=all&_r=0

[26] M. Back *et al.*, "Facebook profiles reflect actual personality, not self-idealization," *Psychol. Sci.*, vol. 21, no. 3, 2010, DOI: 10.1177/0956797609360756.

[27] D. Evans, S. Gosling, and A. Carroll, "What elements of an online social networking profile predict target-rater agreement in personality impressions," in *Proc. Conf. Weblogs Social Media*, 2008, pp. 45–50.

[28] S. Gosling, A. Augustine, S. Vazire, N. Holtzman, and S. Gaddis, "Manifestations of personality in online social networks: Self-reported Facebook-related behaviors and observable profile information," *Cyberpsychol. Behav. Social Netw.*, vol. 14, no. 9, pp. 483–488, 2011.

[29] P. Rentfrow *et al.*, "Divided we stand: Three psychological regions of the united states and their political, economic, social, health correlates," *J. Personality Social Psychol.*, vol. 105, pp. 996–1012, 2013.

[30] D. Quercia, R. Lambiotte, M. Kosinski, D. Stillwell, and J. Crowcroft, "The personality of popular Facebook users," in *Proc. Conf. Comput. Supported Cooperative Work*, 2012, pp. 955–964.

[31] D. Quercia, M. Bodaghi, and J. Crowcroft, "Loosing friends on Facebook," in *Proc. Web Science Conf.*, 2012, pp. 251–254.

[32] D. Quercia *et al.*, "Facebook and privacy: The balancing act of personality, gender, relationship currency," in *Proc. Conf. Weblogs Social Media*, 2012, pp. 306–313.

[33] Q. He, C. Glas, M. Kosinski, D. Stillwell, and B. Veldkamp, "Predicting self-monitoring skills using textual posts on Facebook," *Comput. Human Behav.*, vol. 33, pp. 69–78, 2014.

[34] B. Bi, M. Shokouhi, M. Kosinski, and T. Graepel, "Inferring the demographics of search users," in *Proc. Int. World-Wide Web Conf.*, 2013, pp. 131–140.

[35] J. Wang, S. Faridani, and P. Ipeirotis, "Estimating completion time for crowdsourced tasks using survival analysis models," in *Proc. Workshop Crowdsourcing Search Data Mining/ACM Int. Conf. Web Search Data Mining*, M. Lease, V. Carvalho, and E. Yilmaz, Eds., Hong Kong, China, Feb. 2011, pp. 31–34.

[36] P. Rentfrow *et al.*, "The song remains the same: A replication and extension of the music model," *Music Percept.*, vol. 30, no. 2, pp. 161–185, 2012.

[37] V. Mahalingam, D. Stillwell, M. Kosinski, J. Rust, and A. Kogan, "Who can wait for the future? A personality perspective," *Social Psychol. Personality Sci.*, 2013, DOI: 10.1177/1948550613515007.

[38] C. Jernigan and B. F. T. Mistree, "Gaydar: Facebook friendships expose sexual orientation," *First Monday*, vol. 14, no. 10, 2009. [Online]. Available: http://firstmonday.org/ojs/index.php/fm/rt/printerFriendly/2611/2302

[39] L. Backstrom and J. Kleinberg, "Romantic partnerships and the dispersion of social ties: A network analysis of relationship status on Facebook," in *Proc. 17th ACM Conf. Comput. Supported Cooperative Work Social Comput.*, 2014, pp. 831–841.

[40] C. Anderson, O. John, D. Keltner, and A. Kring, "Who attains social status? Effects of personality and physical attractiveness in social groups," *J. Personality Social Psychol.*, vol. 81, no. 1, pp. 116–132, 2001.

[41] R. Swickert, C. Rosentreter, J. Hittner, and J. Mushrush, "Extraversion, social support processes, stress," *Personality Individual Differences*, vol. 32, no. 5, pp. 877–891, 2002.

[42] A. Friggeri, R. Lambiotte, M. Kosinski, and E. Fleury, "Psychological aspects of social communities," in *Proc. Int. Conf. Privacy Security Risk Trust/Int. Conf. Social Comput.*, 2012, pp. 195–202.

[43] M. Kosinski, "Measurement and prediction of individual and group differences in the digital environment," Ph.D. dissertation, Dept. Psychol., Cambridge Univ., Cambridge, U.K., 2014.

[44] S. Gosling, S. Gaddis, and S. Vazire, "Personality impressions based on facebook profiles," in *Proc. Int. Conf. Weblogs Social Media*, 2007, vol. 7, pp. 1–4.

[45] A. Kramer and K. Rodden, "Word usage and posting behaviors: Modeling blogs with unobtrusive data collection methods," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, 2008, pp. 1125–1128.

[46] S. Vazire and S. D. Gosling, "E-perceptions: Personality impressions based on personal websites," *J. Personality Social Psychol.*, vol. 87, pp. 123–132, 2004.

[47] N. Lane *et al.*, "Enabling large-scale human activity inference on smartphones using community similarity networks (CSN)," in *Proc. 13th Int. Conf. Ubiquitous Comput.*, 2011, pp. 355–364.

[48] H. Lu *et al.*, "Stresssense: Detecting stress in unconstrained acoustic environments using smartphones," in *Proc. ACM Conf. Ubiquitous Comput.*, 2012, pp. 351–360.

[49] K. K. Rachuri *et al.*, "Emotionsense: A mobile phones based adaptive platform for experimental social psychology research," in *Proc. 12th ACM Int. Conf. Ubiquitous Comput.*, 2010, pp. 281–290.

[50] J.-P. Onnela *et al.*, "Structure and tie strengths in mobile communication networks," *Proc. Nat. Acad. Sci.*, vol. 104, no. 18, pp. 7332–7336, 2007.

[51] R. Lambiotte *et al.*, "Geographical dispersal of mobile communication networks," *Physica A, Stat. Mech. Appl.*, vol. 387, no. 21, pp. 5317–5325, 2008.

[52] M. Gonzalez, C. Hidalgo, and A.-L. Barabasi, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.

[53] Y.-A. de Montjoye, J. Quoidbach, F. Robic, and A. Pentland, "Predicting personality using novel mobile phone-based metrics," *Social Computing, Behavioral-Cultural Modeling and Prediction*, vol. 7812, Berlin, Germany: Springer-Verlag, 2013, pp. 48–55.

[54] C. Nass and K. M. Lee, "Does computer-generated speech manifest personality? An experimental test of similarity-attraction," *J. Exp. Psychol.*, vol. 7, no. 3, pp. 171–181, 2000.

ABOUT THE AUTHORS

Renaud Lambiotte received the Ph.D. degree in theoretical physics from the Université Libre de Bruxelles, Brussels, Belgium, in 2004.

He is a Professor in the Department of Mathematics, University of Namur, Namur, Belgium. He was a Research Associate at the École normale supérieure de Lyon (ENS Lyon), Lyon, France; Université de Liége, Liége, Belgium; Université catholique de Louvain, Louvain-la-Neuve, Belgium; and Imperial College London, London, U.K. His research interests include network science, data mining, stochastic processes, social dynamics, and neuroimaging.

Michal Kosinski received the Ph.D. degree in psychology and computer science from the University of Cambridge, Cambridge, U.K., in 2014.

He is a Research Associate at the Computer Science Department, Stanford University, Stanford, CA, USA and the Deputy Director of the Psychometrics Centre, University of Cambridge. He studies big social data and its consequences for privacy, occupational markets, and wellbeing. He also coordinates the myPersonality project, which involves global collaboration between over 150 researchers analyzing a sample of over eight million Facebook users.

INVITED PAPER

Design of Randomized Experiments in Networks

This paper presents novel networked treatment designs and discusses their potential for future research.

By Dylan Walker and Lev Muchnik

ABSTRACT | Over the last decade, the emergence of pervasive online and digitally enabled environments has created a rich source of detailed data on human behavior. Yet, the promise of big data has recently come under fire for its inability to separate correlation from causation—to derive actionable insights and yield effective policies. Fortunately, the same online platforms on which we interact on a day-to-day basis permit experimentation at large scales, ushering in a new movement toward big experiments. Randomized controlled trials are the heart of the scientific method and when designed correctly provide clean causal inferences that are robust and reproducible. However, the realization that our world is highly connected and that behavioral and economic outcomes at the individual and population level depend upon this connectivity challenges the very principles of experimental design. The proper design and analysis of experiments in networks is, therefore, critically important. In this work, we categorize and review the emerging strategies to design and analyze experiments in networks and discuss their strengths and weaknesses.

KEYWORDS | Behavioral science; general; science; sociology; systems, man, and cybernetics

As our day-to-day activities become increasingly embedded in online and digitally enabled environments, the availability of massive scale yet highly granular data on individuals and social interaction enables new avenues of scientific discovery. The promise of big data [1], [2] seems immense—not just for its scale and scope, but perhaps more importantly because highly detailed individual-level data at scale suggest tailored policies that resist reversion to the mean in domains ranging from medicine and public health [3]–[5] to politics, web search [6], business [7], e-commerce [8], and product design [9]. Yet, the promise of big data has recently come under fire for its inability to separate correlation from causation—to derive actionable insights and yield effective policies [10], [11]. This criticism unveils the perhaps lesser known but burgeoning movement of big experiments that is rapidly gaining traction within both academic research and industry practice. The gold standard of causal inference through experimentation is well established in both public and private sectors [12]–[14]. Yet, the realization that our world is highly connected and that behavioral and economic outcomes at the individual and population level depend upon this connectivity challenges the principles of experimental design that lie at the very heart of the scientific process.

Traditional experimental designs that randomly assign populations to control and treatment groups to measure the comparative outcome of a treatment do not account for the networked environment in which we live—the natural connections between subjects in these populations. When the impact of treatment can propagate along these connections, the traditional notions of experimental design break down. It is perhaps not surprising that this realization has chiefly emerged from the blossoming interdisciplinary field of computational social science [15], where the focus of study is on social behaviors that are, by their nature, interactive. Yet the implications of connectivity on experimental design are far reaching and necessarily affect scientific inquiry in multiple domains, including medicine, public health, media, politics, business, biology, epidemiology, sociology, and many others.

However, the natural connectivity of our world does not only present a challenge to the conventional paradigm of experimental design, but also reveals opportunities to

Manuscript received August 8, 2014; revised October 14, 2014; accepted October 14, 2014. Date of publication November 4, 2014; date of current version November 18, 2014.

D. Walker is with the School of Management, Boston University, Boston, MA 02215 USA (e-mail: dtwalker@bu.edu).

L. Muchnik is with the School of Business Administration, Hebrew University of Jerusalem, Jerusalem 91905, Israel (e-mail: lev.muchnik@huji.ac.il).

Digital Object Identifier: 10.1109/JPROC.2014.2363674

leverage connectivity through the creation of novel treatments that incorporate both experimental subjects and the connections between them. Done correctly, networked treatments can allow us to understand the basic dynamics of contagious phenomena that have been found to play a critical role in individual and population level outcomes, such as: the effect of dosage or multiple exposures on individuals and populations [16]–[19], the decay of spreading behavioral and economic outcomes across social distances [20], and the impact of heterogeneity in individual and relationship characteristics on spreading [21]–[28]. In turn, such an understanding will allow us to assess and compare policies designed to promote positive contagions and contain or discourage negative contagions.

In this work, we consider several aspects of networked randomized trial design from the perspective of the experimental setting, the process being studied, and the impact of connectivity. We further address emerging methods to analyze and draw statistical inferences from networked randomized trials. Finally, we present several categories of novel networked treatment designs and discuss their potential to future research.

I. THE IMPACT OF SETTING AND PROCESS ON NETWORKED RANDOMIZED TRIAL DESIGN

Over the past several years, the use of randomized controlled trials in networked environments has increasingly been employed by researchers across a variety of disciplines. Though these works share in common the feature of highly interconnected environments in which they take place, they differ significantly in both intent and approach. Networked randomized controlled trials (NRCTs) can be classified along two dimensions: the setting in which they are conducted and the process they are designed to investigate. First, we consider the setting in which an NRCT is conducted, which has a number of implications on aspects of experimental design, relationship to the networked environment, and on interpretability or generalizability of findings.

A. Setting

There are three primary settings in which NRCTs can be conducted: offline laboratories, online laboratories, and field experiments in real-world settings (often referred to as experiments "in the wild"). The main differences between these settings are: the extent to which the experimenter can control her subjects and the context, the extent to which subjects are aware that they are participating in an experiment, whether networked environments are artificially imposed or organic, the potential scale of the experiment (in terms of population size and experiment duration), the ability to run repeated experiments, the ability to recruit and maintain subject participation, and the amount and type of information on subjects that is available for posterior analysis.

Offline laboratory settings have traditionally been used in the fields of psychology, economics, and sociology [29], [30]. In this setting, participants are typically recruited, invited into a highly controlled physical environment, and given instructions on how to participate across multiple phases of the experiment according to well-established protocols. Offline lab settings offer the advantages of strict experimenter controls (over the conditions of the environment itself, constraints on subject behavior, and the nature of subject interactions and information flow). For example, experimental subjects can be deliberately primed, exposed to controlled situations, given background, and even instructed to act or interact with one another in a particular manner. The advantage of strict control, however, is often accompanied by the important tradeoff that subjects are explicitly and constantly aware of their role as experiment participants, and this awareness may cause them to act, react, and interact differently from their natural behavior in organic environments and in cases where they do not believe their behavior is being observed and assessed. This limitation may have important implications on the generalizability and applicability of findings to policy considerations [31]. Beyond aspects of control, subject recruitment is often limited by proximity to the lab, time availability, and effectiveness of recruitment incentives. The limitations of subject recruitment have two important implications: First, it constrains the demographics of experimental populations (and thereby the generalizability of findings) and second, the overall population size and duration of the experiment. In addition, the research questions that can be addressed by the experiment are frequently limited by the premises at which it is set. Offline lab settings also have advantages and disadvantages with regard to the networked environment itself. In these settings, researchers may completely specify network connections between subjects and control communication or other forms of interaction along these links. However, networks imposed by researchers may be very different in structure from organic networks, and artificially imposed network links may lack real social context, potentially making them a poor proxy for the real social environments they may intend to represent. These conditions facilitate investigation of well-defined situations, such as a collaborative solution to the network coloring problem [32], [33], convergence to consensus through biased voting [34], or the impact of network structure on the performance of prediction markets [35], [36].

Online laboratory settings are relatively recent and primarily facilitated by the pervasiveness of online technologies and the emergence of online social network platforms and microlabor markets (such as Amazon's Mechanical Turk [37]). These settings replicate the spirit of the offline lab in that subjects are explicitly aware that they are participating in an experiment, may be primed, given background information, and requested to act or interact with one another in a particular manner. To some extent, online lab settings reduce constraints on experimental scale and subject recruitment in terms of geographic proximity and duration of participation. Importantly, unlike their offline counterparts, online lab settings can leverage existing platforms to enable subject recruitment at significantly reduced costs [38], [39] and thus have the potential to enable experimentation at much larger scales, though, like their offline counterparts, online lab settings may also suffer from concerns of generalizability arising from the makeup of microlabor markets employed for subject recruitment [40]. In addition, these settings can also leverage the application programming interface (API) of existing platforms or data sharing agreements with their operators to collect detailed information about subjects, their social network connections, and to control or mediate subject interactions [41]–[44]. However, these environments necessarily sacrifice strict experimental control in terms of the conditions of the offline environment itself, constraints on subject behavior, and the nature of subjects (potentially unrecorded) offline and online actions and interactions, as well as information flow to and (in some cases) between subjects. Experiments in online lab settings also face a number of new challenges such as maintenance of subject participation (e.g., user churn) and concurrency of subject participation (i.e., experiment design may require simultaneous presence of the subjects)[1] [40]. As in the case of offline labs, the findings and inferences from experiments conducted in online lab settings may have limited applicability to real-world environments because individual behavior may be affected by the knowledge that subjects are part of an experiment and are being observed. Unlike offline lab settings, online labs that leverage existing social network platforms permit experiments in real networked environment while exerting some degree of control of interactions and information flow along network links (e.g., [44]). Thus, online lab settings circumvent some limitations of their offline counterparts making them uniquely suited to address well-defined situations such as the role of network in cooperation [39], [45], public goods [39], and investment games [44] as well as its impact on health behavior [41]–[43].

In contrast to offline and online labs, field experiments in real-world settings do not exert strong controls over subjects' environments, but instead assess the impact of randomized assignment directly in the natural environment of the system being studied [46]. Online field experiments in particular can provide researchers with detailed data on subject behavior (online and even offline)[2] that is not biased by knowledge of participation in the experiment[3] and can be conducted at extraordinarily large scales and over arbitrarily long durations. In some sense, online field experiments are a natural extension of A/B testing procedures that have become part of the standard policy for large online platforms to assess and evaluate features or the impact of platform design elements on the overall user experience [13], [47]–[49]. Because these settings facilitate experiments that can be conducted without or with limited subject knowledge, care must be taken to assess the ethical considerations of these practices and to abide by the standards of practice governing human subjects research. This concern tends to be more central to experiments addressing fundamental social science or economics research questions than in the case of routine A/B testing. Controversy surrounding recent research employing an online field experiment to study emotional contagion on Facebook emphasizes these concerns [50], [51]. In addition, researchers that conduct experiments in real-world networked environments with treatment impacts that can propagate should also consider the ethical implications of treatment impact on individuals outside the experimental population. It should be noted that despite the necessity for strong ethics, field experiments in real-world settings provide strong inferences and insights directly applicable to real-world systems and thus play a critical role in assessing the potential efficacy of important social and economic policies. Additional aspects of design of field experiments in natural settings relate to the concerns that the desired interventions should appear to be organic, in some cases not clearly detectable between subjects, and generally should not observably interfere with the normal operation of the community, platform, or online system. These concerns are important for rigorous experimental design but also to assure that experimental interventions do not adversely affect the business of firms that collaborate with researchers. Like online lab settings, online field experiments are limited by the capacity for experimenters to design interventions or otherwise control the environment. For example, it may be more difficult to expose subjects to a desired intervention, and other experimental controls may be limited by the features of the online platform or system. In many cases, field experiments identify effects of specific platform features, such as the impact of (in)visibility of user activity on peer interactions on an online dating site [52], the impact of social cues in word-of-mouth advertising on ad performance [53], the role of social platforms in diffusion of information [54], the study of the mechanism of coupon sharing on Facebook [55], and the value of the content author's identity in evaluation of that content by the reader on news aggregation websites [56]. Besides the advantage of conducting the experiment in organic

[1] See, for example, the sections on dropouts and the waiting room in [39].

[2] For example, many online platforms are location aware.

[3] In many cases, experiments conducted in natural settings may notify subjects indirectly through posted policies on user research in the platform or online site's terms of use.

settings and thus observing natural behavior, real-world settings often enable research at immense scale, facilitating observation of subtle effects or heterogeneous response to interventions.

Finally, while not the focus of this work, the occurrence of natural experiments arising from exogenous variation in real-world systems provides yet another setting in which researchers may pursue causal inference. Unlike formal experiments which require significant investments of time and resources from experimenters, occasionally induce ethical concerns, and may noticeably interfere with user experience, the potentially widespread occurrence of natural experiments may permit causal inference on observational data at large scales and may be used prior to experimentation to inform experiment design. Having discussed the implications of setting on the design of NRCTs we now turn to a discussion of the process that NRCTs are designed to investigate.

B. Classification by Process

Experiments in networked environments can also be classified by the process they are designed to investigate, including the exploration of social and economic behaviors, the underlying dynamic microscopic and macroscopic mechanisms governing these behaviors, and the resulting dynamics of outcomes at individual, group, and population levels. Many NRCTs focus on investigation of propagation processes such as dissemination of innovation, spread of information and behaviors, or adoption of new products. Identification of factors affecting these processes is vital for informing managerial or public policies intended to promote or discourage population-level outcomes. Factors that affect process dynamics include initial conditions (such as targeting or seeding); dosage and temporal aspects (such as the extent and timing of multiple exposures); the willingness of subjects to contribute, prevent, or direct the viral spread; the susceptibility of subjects to peer influence, the social network topology; and modification of the process itself (e.g., viral product design [22]) In practice, policies may need to leverage one or more of these mechanisms to achieve a desired outcome [57].

The process under investigation is often signified by how the experimenter measures the response to their intervention. When the effect of intervention(s) can propagate, the experimenter is not limited to analyzing the response behavior of directly treated subjects, but may instead focus on the response behavior of other subjects or groups in the population (such as peers of directly treated subjects or groups of locally connected treated subjects). Analysis may therefore focus on one of three aspects: the direct effect of the treatment on the treated (ETT) subjects, the effect of the treatment on the cotreated (ETC), or the effect of the treatment on the untreated (ETU).[4] To avoid ambiguity, we adopt a simple definition of treatment that is defined for each experimental subject as the alteration of that subject's experimentally controlled experience. We leave discussion of more complex networked treatments not covered by this definition to Section III. The effect of treatment may be measured at the level of individual subjects, aggregated over groups of subjects, or aggregated effects of the treatment on the population at large. In addition, researchers may be interested in how the effect is moderated by individual attributes, local-network attributes, or the global structure of the network. In this section, we categorize existing research by process and discuss the implications on experimental design and choice of setting.

Networked experiments that study processes concerned with the effect of treatment on the treated (ETT) represent the extension of conventional nonnetworked experiments to networked environments. It is important to note that for many processes of interest (particularly those that involve social components) traditional experimentation may be affected by an underlying network, even when the network is not explicitly observed or recognized by experimenters. In some cases, interaction between subjects may be an unavoidable nuisance, while in other cases, it may be central to the process under investigation. For example, Bapna *et al.* [52] study the impact of enabling anonymous profile viewing for users of an online dating site, an intervention that is meaningless in the absence of social interaction. In another experiment Bakshy *et al.* [53] vary the number and the intensity of social cues accompanying online ads to establish the degree to which they can affect the ad performance. One distinct class of research questions that focus on ETT addresses subjects' response to population level social signals such as conformation to peer pressure. First identified with the now classical sociologist techniques in lab or small-scale field experiments conducted in the 1950s and 1960s by Asch [58] and Milgram *et al.* [59], these phenomena can now be examined at scale. For example, in a sequence of experiments, Salganik *et al.* [38], [60] study the impact of popularity-based content ordering on the propensity to consume cultural products (music). In these experiments, the authors randomize perceived popularity of songs to distinguish the impact of popularity on subjects' decisions to consume music from that of song quality. The rising prevalence of ranking and rating mechanisms in virtually every domain makes these types of experiments both theoretically and practically important. The sheer scale of the data can permit subtle inferences that require high sensitivity and provide enough

[4] This terminology should not be confused with the traditional terminology of the average treatment effect on the treated (ATET) and the average treatment effect on the untreated (ATEU) which pertain to analysis in nonnetworked environment to provide counterfactual estimates that avoid selection bias in the designation of treated populations.

resolution to understand the moderating effect of different personal and content attributes on consumption decisions. Depending on the setting and control available to experimenters, connectivity between subjects may allow for undesirable spillover effects (interference) that contaminate or bias inferences on ETT. The potential scale and the scope of networked experiments can facilitate inferences on the ETT that emerge as a result of feedback mechanisms (due to spillover effects). For example, individuals subject to treatment can indirectly influence others by contributing their manipulated opinion to population-level social signals. Such spillover effects may affect both untreated and treated subjects (through feedback) resulting in herdlike phenomena that has the potential to sway collective (population-level) behavior, potentially in undesirable ways, leading to important implications for rating, ranking, and collaborative feedback systems. This effect was demonstrated at scale by Muchnik *et al.* [61] who show that minor manipulation of the perceived scores of user-generated comments guided consequent user votes and resulted in herding, significantly affecting the content's final score. Subsequent research by Godinho de Matos *et al.* [62] found that manipulation of rank and population-level social signals for video-on-demand titles yields only a short-lived effect on herding behaviors, emphasizing that such signals may be highly context dependent. From the experiment design perspective, the feedback following the randomized manipulation of the content or its ranking may mingle the treatment with endogenous processes. In fact, due to the aggregation of the collective opinion into population level social signals, all but the first impression following the treatment are conditional on the response (or nonresponse) of the preceding subjects. We discuss detailed strategies to address the related interference issues in Section II.

Networked experiments that study processes concerned with the effect of treatment on the cotreated (ETC) include processes that involve local network externalities. Understanding such processes is central to explaining the value of network goods, products, or features and necessary for modeling of propagation of knowledge, rumors, and information in general. The recent emergence of pervasive online social platforms enables experimentation on ETC-related processes that can yield relevant insights of value to both platform owners and academics. For example, recent work on network bucket testing extends A/B testing procedures to assess the ETC of social product features. Additionally, many platform sponsors assess social features through beta rollouts (e.g., Gmail) that allow users to invite their peers to coadopt, making inferences on ETC of high practical importance. For this reason, studies of ETC are often conducted in real-world settings, though the available controls of offline and online lab settings make them equally suitable to study cotreatment.

Processes concerned with the effect of treatment on the untreated (ETU) are the focus of the rapidly expanding field of research into contagious phenomena across multiple disciplines. Many recent randomized experiments conducted in networks examine contagion processes in the context of diffusion of behaviors (e.g., voting [20] or health behavior [41]), emotions [50], peer influence, or product placement (seeding) [21], [63], [64], which aspire to inform polices aimed at the promotion or containment of contagions in social networks. These studies are designed to identify the impact of a variety of factors on contagious spreading. Several studies investigate the moderating effect of individual characteristics [52], dyadic properties [21], [53], [64] as well as the impact of attributes of a spreading product, norm, or information [20], [22], [50] on diffusion processes on networks. Causal identification of factors that affect subject behavior can be achieved through exogenous manipulation of these factors, allowing researchers to distinguish causal impact from alternative explanations of correlated behavior such as homophily, assortative mixing, and other endogenous confounds [65]–[67]. These experiments are based on selective application of treatments to focal subjects and observation of the response of their immediate or remote peers. Typical treatments include randomized gifting, variation of pricing, or manipulation of product features. More sophisticated treatments focus on randomly controlling the interaction between individuals and their peers, aiming to test how peer influence is moderated by subject, peer and dyadic characteristics. For example, Aral and Walker [64] test the moderating effects of individual and dyadic characteristics on word of mouth by issuing Facebook notifications to randomly chosen peers of experimental subjects. Such networked treatments go beyond the definition of simple treatment that we have adopted here. We discuss these types of treatments in more detail in Section III. Other contagion experiments examine the effect of local and large-scale network topology on diffusion of information and behaviors [32]–[34], [39]. The general goal of these studies is to detect the effect of network attributes (such as degree, clustering, assortativity) on network diffusion processes and the effect they have on individual and collective behavior such as convergence to consensus [34], [39], a collaborative solution to network coloring problems [32], [33], and the spread of health-related behaviors [41]–[43]. Experiments on contagion processes may constrain the choice of experimental setting. For instance, exogenous manipulation of local or global social network structures requires a setting where tight control over individuals' connections is possible. Such control can be achieved in offline or online laboratory settings where the experimenter has full control over connections and/or information visibility, but may be less feasible in real-world settings where connections emerge organically and cannot be exogenously manipulated[5] [32]–[34], [39], [41], [42]. On the other hand, studies that examine the (potentially subtle) impact of individual attributes on contagions typically require rich data sets at large scales that cover the wide spectrum of attributes. Such studies are, therefore, best suited to settings that enable large-scale experimentation where data is readily available, as is the case for real-world field experiments conducted on online platforms.

II. IMPACT OF CONNECTIVITY ON RANDOMIZED TRIAL DESIGN

The consequence of connectivity on inference in randomized trials is best understood by examining the Rubin causal model, which presents a fundamental approach to drawing causal statistical inferences from randomized experiments. A chief assumption of this approach is the stable unit treatment value assumption (SUTVA) which demands that the observed outcome on one unit (subject) should not depend upon treatment assignments to other units (subjects) [68], [69]. When the effect of treatment can propagate, this assumption is violated and the standard machinery of statistical inference from randomized trials must be reevaluated. In some cases, where propagation of treatment effects is well understood, the SUTVA can be reestablished by redefining treatment to multiple treatment specifications that include indirect spillovers. However, in the highly connected environments in which we are interested (and particularly where the nature of treatment propagation is unknown), simple respecification of treatment to rescue the SUTVA is not feasible.

In recent work, Manski [70] has taken the first steps toward building a theory of identification in the presence of interference by extending the SUTVA from the classical assumption of noninteracting units, which he refers to as the assumption of individualistic treatment response (ITR), to define multiple classes of assumptions based on the nature of interaction (or lack of interaction) between units. For example, he defines the assumption of constant treatment response (CTR) as the case when each individual in the experimental population has some reference group (of other units or subjects) for which his or her outcome remains constant when treatment varies beyond his or her reference group. He further relates these assumptions to models of endogenous interactions through systems of simultaneous equations that connect treatment and outcomes of all individuals in the population to the outcome of any particular individual. These considerations lead to restrictions on when inference from observed outcomes can be point identified and, importantly, how this relates to treatment designation.

[5]We note that online lab settings and real-world online settings that utilize platform features may effectively alter network structures exogenously by disabling certain types of interactions, rendering these settings suitable for studies aimed at inferring the impact of local or global network structures on contagion dynamics.

Practical strategies to account for connectivity in randomized trials are currently an active topic of research and fall into two general categories: inference strategies and design strategies [64]. The former strategies address interference after an experiment has been conducted, during the inference or analysis phase, while the latter strategies address the potential for interference prior to experimentation by modifying aspects of the design of randomized trials, such as treatment assignment procedures, to minimize interference.

To clarify our discussion of these strategies, we introduce some terminology to describe treatment and exposure to treatment. For the purposes of simplicity, we assume for now that experimental treatments apply directly to individuals (or units) in the population and leave complex treatment types that may include simultaneous experimental controls on individuals, their peers, and the nature of their interaction(s), for subsequent discussion. We also assume, for simplicity of discussion, that treatments are temporally static, assigned prior to the experimental period and consist of only one kind of treatment (i.e., treatment or control; though these definitions may easily be extended to the case of multiple treatment types). We define direct treatment as the alteration of each individual's experimentally controlled experience throughout the course of the experiment, as specified by the direct treatment vector T_i^{dir}, where i indexes experimental subjects. This follows the conventional usage of the term "treatment" in traditional RCTs, and its assignment is directly controlled by the experimenter. In contrast to direct treatment, we define indirect treatment as the experience induced on peers of directly treated users (through their direct connection or through one or more pathways of multiple connections in the network) as a consequence of direct treatment, as specified by the indirect treatment vector T_{ik}^{ind}. Unlike direct treatment, indirect treatment is not exogenously assigned, but arises instead from both direct assignment, the (often endogenous) network itself, and the (often endogenous and unknown) dynamics of propagation of the impact of direct treatment. The subscript k is included to enumerate the multiple types of indirect treatment that arise through exposure even given our assumption of one kind of direct treatment. For example, one type of indirect treatment may be defined as having one and only one treated neighbor (regardless of treatment assignments at larger network distances from the subject); another type may be defined as having two treated neighbors who are not connected to one another (regardless of treatment assignments at larger network distances from the subject). As these examples suggest, the multiplicity of indirect treatment types depends on assumptions about exposure and propagation. As a consequence, indirect treatment may also be time dependent. For completeness, we also define the total effective treatment as the combination of both direct and indirect treatment T_k^{tot}.

A. Inference Strategies

Inference strategies attempt to remove or reduce bias and/or variance from estimates that identify the impact of treatment in connected settings and typically assume a specified vector of treatments. Inference strategies are distinguished by the type of estimation strategy, from the fundamental estimate of the average impact of treatment to more sophisticated modeling techniques.

In some recent work, researchers have developed methodologies to partially account for statistical interference in NRCTs through modified average treatment effect (ATE) estimators with reduced bias [71], [72]. In these methodologies, an exposure model is assumed and employed to enumerate multiple total treatment types, T_{ik}^{tot}. A modified Horvitz–Thompson or Hajek estimator is then constructed to account for the bias introduced by the propensity to receive any of the total effective treatment types. The multiplicity of the total effective treatments is determined by assumptions of the exposure model. For example, for an exposure model that assumes propagation to fall to zero beyond one network link, all individuals with no treated neighbors will have a total effective treatment equal to their direct treatment, regardless of the treatment status of peers at network distance greater than one. However, it is clear that for arbitrary exposure models (where the propagation of the treatment effect may not fall sharply with increasing network distance), estimating the causal impact of the treatment becomes severely limited, as the number of potential indirect treatments (K) becomes increasingly large relative to the size of the experimental population. Coppock and Sircar summarize this difficulty succinctly:

> "The basic difficulty inherent in design of experiments facing interference between units is that it reduces power. If units are exposed to complex spillovers, the outcomes revealed by those units are not useful for the estimation of any quantities of interest" [73].

Nonetheless, modified estimator approaches may be particularly fruitful when strong assumptions of limited propagation apply or when the experimenter can exert strict control over propagation. However, when knowledge of the propagation (and hence exposure) is unknown, practitioners must turn to empirical evidence to first adjudicate between multiple potential exposure models. It is important to note that the statistical interference methods discussed above are not designed to discriminate between different exposure models. This highlights a critical challenge in analysis of networked experimentation in novel contexts: researchers must simultaneously estimate both the treatment impact and the nature of exposure dynamics.

Other inference strategies go beyond modification of estimates of the ATE impact and incorporate constraints on inference in more sophisticated approaches to model treatment impact. Modeling in NRCTs has three primary advantages over ATE estimation. First, use of models that incorporate interactions of characteristics or attributes with both direct and indirect treatments allow inferences surrounding the heterogeneity of treatment impact. Such inference can be used to understand and predict how different subpopulations would respond to treatment.[6] This is particularly important from the standpoint of personalized policy development. While true assessment of the efficacy of personalized policies should be verified by evaluating interventions specifically designed to affect targeted subpopulations, inferences on heterogeneous treatment impact can act as a guide to develop personalized policies by identifying subpopulations (from the wide range of possibilities) for which treatment impacts significantly differ. Second, modeling permits identification of moderators of treatment impact *ceteris paribus*, allowing researchers to partially disentangle the treatment impact of correlated characteristics, provided there is significant diversity in subject populations. Third, modeling strategies allow researchers to employ tools such as censoring, stratification, and matching to estimate the impact of indirect treatment on individuals who have received different exposures relative to those in appropriate reference groups that have not. In models that employ duration analysis, censoring techniques can be used to censor outcomes of users only after they are exposed to complex indirect exposures. This technique allows researchers to reduce bias in estimates of treatment impact while both retaining the maximal amount of outcome data in their analysis and correctly parameterizing their ignorance regarding what might have happened had complex indirect exposure not occurred. For example, Aral and Walker [22] employ censoring in hazard models to exclude subject outcomes from analysis only after they have been indirectly exposed to multiple treated peers (with potentially different treatments). Unlike the modifications to ATE estimators discussed above, censoring techniques do not require complete specification of an exposure model, but instead assert limiting assumptions regarding exposure in exchange for both a loss of statistical power for censored observations and the inability to estimate the impact of some complex exposures. This is an important tradeoff. Stratification (in nonduration modeling) and dynamic risk group assignment (in duration modeling) further allow researchers to partition subjects according to different indirect exposures they may have received and separately estimate the impact of these indirect exposure on subject outcomes. Stratification on indirect exposure types is also subject to assumptions regarding the nature of exposure (as indirect exposure

[6] Recent work on analyzing heterogeneous treatment effects with dependent data provide a variety of bootstrap methods to properly handle uncertainty [85].

types must be specified prior to stratification), but also does not require a complete specification of the exposure model. Instead, complex exposure types excluded from any strata or risk group are effectively censored (in duration models) or truncated (in nonduration models), sharing the associated tradeoffs with censoring discussed above. Importantly, because indirect exposure is endogenously determined (by the natural connectivity of the network and in some NRCTs potentially also by endogenous propagation of the treatment), concerns of generalizability apply. Specifically, populations receiving different types of indirect exposure to treatment may be fundamentally different (in terms of observable and unobservable/latent characteristics) from the representative population at large. Researchers employing these techniques should, therefore, take care in generalizing inferences on the impact of complex indirect exposures to situations that would not arise organically (such as policies that exogenously determine complex indirect exposures). Matching techniques generally may be employed to establish appropriate reference groups and specifically to control for the propensity to receiving a particular type of indirect exposure to treatment and to balance the makeup of direct or indirectly treated populations relative to controls. Matching on propensity to receive a type of indirect treatment is comparable to the modifications to ATE estimators discussed above, but matching techniques can be generalized to simultaneously account for both propensity to be exposed and endogenous variations in the makeup of subpopulations that receive different types of direct and indirect exposure. It is important to note that relative to fundamental ATE approaches, modeling approaches may often involve strong assumptions regarding the mechanism of treatment response. For this reason, researchers must establish that these assumptions are theoretically grounded and reasonable through empirical validation and ensure that the robustness of inferences to model specification is thoroughly explored.

B. Design Strategies

In contrast to inference strategies, design strategies alter aspects of the design of the experiment itself in order to constrain the manner of interference between subjects. Typically design strategies involve rearranging assignment of treatment to subjects in a manner that incorporates information on network connectivity. Existing design strategies fall into two categories: treatment clustering strategies and treatment separating strategies. Treatment clustering strategies seek to closely approximate the counterfactual conditions in which the entire network is exposed to either treatment or control by assigning subjects in well-defined local subnetworks the same treatment. In contrast, treatment separating strategies seek to assign treatments to experimental subjects that are well separated from one another in network distance in order to minimize interference. Existing design strategies to deal with interference also differ by whether they are appropriate for making unbiased inferences on the effect of treatment on the treated (ETT), cotreated (ETC), or untreated (ETU) members of the populations and the extent to which they are suitable for empirically inferring (rather than assuming) exposure dynamics.

Treatment clustering strategies relate treatment designation to the natural structure of the network in terms of clusters, components, and communities.[7] These strategies stem from attempts to extend A/B testing to networked environments where treatment is oriented around enabling new features or products in social network platforms that exhibit strong local network externalities. For example, the evaluation of a new social messaging feature would be inaccurate if the feature was not simultaneously available to individuals and their direct network peers with whom they would typically communicate. Treatment clustering strategies use (a variety of) algorithms to assign the same treatment to clusters of well-connected nodes [17], [74]. Ugander *et al.* [75] use the terminology "network exposed" to describe the condition under which an individual and some sufficient number or fraction of his or her peers have received the same (direct) experimental treatment. They show that, using their technique of graph cluster randomization, an efficient dynamic program can be used to exactly calculate the probability that each individual in the network is network exposed. When an exposure model is specified, these probabilities can be used in modified ATE estimators to reduce bias. Moreover, they also show that under the right conditions, graph cluster randomization can significantly reduce ATE estimator variance. Airoldi *et al.* [76] also consider a simple sequential randomization algorithm that clusters direct treatments in local networks as well as an insulated neighbor randomization algorithm that relaxes treatment conditions to partial neighborhoods to yield a higher probability of valid causal estimates where treatments are matched with counterfactual controls. In subsequent work, Eckles *et al.* [72] point out that many tractable exposure models do not realistically account for the role of peer effects in mediating exposure. Instead, they consider dynamic outcome generating processes in discrete time for which a subject's response at time t depends upon their own direct treatment, as well as the direct treatment and behavior of their peers at time $t-1$. Outcome generating processes go beyond specification of exposure alone and specify a mechanism by which responses are induced by direct and indirect treatments. They employ graph cluster randomization on several artificial models of networks assuming fractional neighborhood treatment response

[7] Clusters are subgraphs into which an overall network is partitioned according to some clustering rule. Components are sets of nodes that are connected to one another via network paths of any length. Communities in networks are defined as sets of nodes which are well connected to one another and relatively sparsely connected to other nodes in the population.

(FNTR) in which a subject is assigned the treatment condition of a specified fraction of their peers. Subjects without a sufficient fraction of peers assigned to treatment or control conditions are excluded from analysis. Using simulations of outcome generating processes on artificial network models, they show that in the presence of peer effects that mediate exposure, graph clustering randomization can reduce bias in modified ATE estimators with comparably small increases in estimator variance when the network itself exhibits sufficient clustering. Thomas and Finegold [77] employ a form of indirect treatment clustering. They consider random treatment assignment and use a pseudorandomized trial (where mock treatment designation does not alter user experience whatsoever) to demonstrate that simple *t*-tests on the impact of indirect treatment (ETU) spuriously bias *p*-values toward zero. They implicitly assume that exposure does not extend beyond a network distance of one and consider permuting direct treatment assignment so that all peers of directly treated subjects have the same unequivocal indirect treatment designation. They show that clustering of indirect treatments restores uniform *p*-value distributions for *t*-tests on the impact of indirect treatment, as would be expected given the mock nature of the treatment. As the above discussion should make clear, treatment clustering strategies can reduce bias and variance in inferences on ETC. However, these strategies necessarily reduce heterogeneity in types of indirect exposure, making them less suitable for inferences on the effect of the treatment on the untreated (ETU), including the ability to empirically evaluate the dynamics of contagious phenomena, such as how multiple indirect exposures add together or how exposure decays over social distance. In some cases, indirect treatment clustering strategies may be appropriate for inferring the effect of treatment on the untreated (ETU), when exposure does not extend beyond a network distance of one. Importantly, treatment clustering strategies may yield unbalanced assignment of nodes to treatment conditions in terms of individual-level or network characteristics of subjects (such as degree). Specifically, Ugander *et al.* [75] point out that subjects with high network degree are less likely to be assigned to extreme definitions of cotreatment (e.g., an effective treatment where most or all of a subject's peers have the same treatment). Likewise, Thomas and Finegold [77], who primarily focus on the impact of indirect treatment, discuss concerns of selection bias for indirectly treated subjects in terms of bias in the distributions of individual characteristics (that may arise from, for example, homophily), and network characteristics (such as degree), that arise as a consequence of designating treatment either randomly or with treatment clustering strategies. While reweighting designation of direct treatment can alleviate selection bias on indirectly treated subpopulations, it necessarily induces selection bias in the directly treated subpopulations, as they point out. One promising approach to address concerns of balance is presented in the recent work by Nishimura and Ugander [78] on graph partitioning.

Treatment separating strategies attempt to reduce interference between subjects by constraining direct treatment assignment to subjects that are well separated from one another. Coppock and Sircar [73] define the SUTVA degree (λ) as the network distance beyond which spillover does not occur.[8] In this methodology, well-defined direct and indirect treatment types on which the experimenter would like to make inferences are specified in advance and all other (complex) exposures to treatment are minimized through a two-stage random direct treatment assignment algorithm that incorporates the assumption of the SUTVA degree. Modifications of the direct treatment assignment algorithm can be performed to permit inferences on the dynamics of contagious phenomena such as how indirect exposures to treatment add together or decay over social network distance. Analysis procedures may also employ modified estimators or other modeling techniques that adjust for the propensity to receive an indirect treatment. Consequently, treatment separating strategies are ideal for estimating the effect of the treatment on the untreated (ETU). Because this strategy primarily seeks to separate treated subjects from one another in network distance, it is less appropriate for inferring the effect of the treatment on the cotreated (ETC) when a substantial number of cotreatments among directly connected individuals is desired. It is important to note that the assumption of a SUTVA degree excludes cases when maximal spillover distance can depend on the number of indirect exposures. For example, in complex contagion scenarios, a subject may be more likely to be affected by multiple directly treated peers at distance $\lambda + 1$ than by a single peer at the same distance. *Post hoc* inferences on how exposure adds and decays over social network distance (within the SUTVA distance) obtained from treatment separating approaches can be examined to evaluate whether this is a concern. Practitioners may wish to modify the direct treatment assignment algorithm to reduce multiple exposures at the cost of reducing treated population sizes and statistical power. Just as treatment clustering schemes may induce selection bias in individual-level or network-level characteristics of directly or indirectly treated populations, as a consequence of clustering, treatment separating schemes may also induce a similar selection bias. The algorithmic removal of subjects within a distance λ from treated and indirectly treated subjects from consideration to receive a direct treatment could impact the balance of treated and indirectly treated subpopulations in terms of individual- and network-level characteristics. The presence of homophily on individual-level characteristics in a variety

[8]The definition of SUTVA distance is closely related to the concept of *r*-nets in metric spaces, which is discussed by Ugander *et al.* [75].

of real-world networks emphasizes this concern. As such, care should be taken to ensure that directly and indirectly treated subpopulations are balanced with respect to one another and any reference groups. When this is not the case, the two-stage random direct treatment assignment algorithm can be modified to reassert balance.

Interestingly, both treatment clustering and treatment separating strategies require assumptions about exposure distance introduced through the choice of the cluster size in the former case, or through the specification of SUTVA degree in the latter case. When empirical evidence is unavailable to inform these decisions, practitioners may employ combinatorial designs to vary cluster sizes in treatment clustering strategies or to empirically infer decay of exposure across social distance in treatment separating strategies. In many circumstances, quasi-experiments that apply matching to observational data may act as a useful guide to inform experimental design surrounding exposure assumptions and the requisite statistical power necessary to infer significant effects [63], [66], [79]. In addition, both treatment clustering and treatment separating strategies assume that the network structure is known. While unbiased sampling can be achieved through a variety of means, e.g., [80] and [81], it may not always be feasible. When only partial information on network structure is available, adaption of the strategies presented here in combination with network sampling techniques may be required. This is another avenue for potential future research.

III. NETWORKED TREATMENTS

The natural connectivity of our world does not only present a challenge to the conventional paradigm of experimental design, but also reveals opportunities to leverage connectivity through the creation of novel treatment mechanisms that incorporate both experimental subjects and the connections between them. Where simple treatments are defined as those that are applied to and alter an individual subject's experience, networked treatments involve interventions that may alter how connected subjects interact with one another, encourage or incentivize a subject to promote or influence the actions of one or more peers in a particular way, affect shared experiences and interactions between groups of subjects, or even encourage the formation of new connections between subjects. Such networked treatments are in part made possible by the emergence of online social networking platforms and other digital social environments that permit firm mediation of social interactions to both platform owners and to other researchers through APIs [49], allowing for varying degrees of experimental control along the channel of social interaction [55], [82]. Networked treatments also enable experiments that can act as important test beds for emerging social policies aimed at producing or altering population-level change. Categories of networked treatments include peer-oriented incentive schemes, communication-altering schemes, subject-grouping schemes, and network topology manipulation schemes. Peer-oriented incentive schemes reward subjects when their peers take a particular action, such as purchasing a product or service [83], making certain choices [34], [39], spreading a particular piece of content or message (such as encouragement to have a flu shot or get an HIV test), or encouraging referral chains [84] that yield desired outcomes (such as a solution to a crowd-sourced problem). Communication-altering schemes may send automated referrals from a subject to his/her peer [22], randomize the target of automated messages from a subject to randomly chosen subsets of his/her peers [21], or even block or moderate information exchange between subjects [50], [83]. Subject-grouping schemes may randomly designate experimental subjects to social environments [such as pairing participants with online health buddies [41]–[43] or designating subjects to online study groups in massive open online courses (MOOCs)] contingent upon subject or environmental characteristics. Network topology manipulation schemes are designed to test the implications of network topology for social computation processes such as collaborative problem solving of competitive games [32]–[34], [39], [45]. Depending on type, instantiation, and context, networked treatments may either remain susceptible to or circumvent interference effects. Future research should evaluate when and to what extent emerging design and inference strategies to address interference can be extended to networked treatments or whether new strategies are required.

IV. DISCUSSION/CONCLUSION

The increasing prevalence of networked environments and the natural connectivity of our world present both challenges to existing design and analysis methods for randomized trials and opportunities to conduct novel experiments involving networked treatments. It is likely that large-scale experimentation in social networks will lead to significant advances in the social sciences, just as conventional randomized controlled trials advanced medicine in the second half of the 20th century. However, just as the widening use of RCTs in medicine, psychology, and other domains necessitated the development of specialized methodologies and analysis techniques, the emergence of NRCTs introduces a number of new challenges, issues, and concerns. While we have systematically reviewed emerging approaches to address these topics, the study of the implications of setting, process, and connectivity on design and analysis of networked randomized trials is still very much in its infancy. Well-designed networked treatments and other novel approaches to the mechanism of randomization [64] may circumvent many of the issues discussed here. Future research employing networked treatment designs should thoroughly consider issues of inference in the presence of interference. More generally, practitioners conducting NRCTs should evaluate the suitability of the design and analysis strategies outlined here to their particular context. The dual challenge of estimating both the impact of experimental interventions that can propagate and the dynamics of propagation itself may call for the development of concurrent design strategies that allow for simultaneous empirical inferences on the former and the latter. The development of analysis techniques that can discriminate between multiple models of propagation or outcome generating processes is also an important avenue for future research. ■

Acknowledgment

The authors would like to thank J. Ugander and the anonymous reviewer for their insightful comments and discussion.

REFERENCES

[1] V. Mayer-Schönberger and K. Cukier, *Big Data: A Revolution That Will Transform How We Live, Work, Think*, 1st ed. Chicago, IL, USA: Eamon Dolan/Houghton Mifflin Harcourt, 2013, p. 256.

[2] F. Provost and T. Fawcett, *Data Science for Business*. Sebastopol, CA, USA: O'Reilly Media, 2013, p. 414.

[3] M. A. Hamburg and F. S. Collins, "The path to personalized medicine," *New England J. Med.*, vol. 363, pp. 301–304, 2010.

[4] T. B. Murdoch and A. S. Detsky, "The inevitable application of big data to health care," *J. Amer. Med. Assoc.*, vol. 309, no. 13, pp. 1351–1352, Apr. 2013.

[5] P. B. Jensen, L. J. Jensen, and S. Brunak, "Mining electronic health records: Towards better research applications and clinical care," *Nature Rev. Genetics*, vol. 13, no. 6, pp. 395–405, Jun. 2012.

[6] P. Brusilovsky, A. Kobsa, and W. Nejdl, *The Adaptive Web: Methods and Strategies of Web Personalization*. New York, NY, USA: Springer-Verlag, 2007, p. 766.

[7] H. Chen and V. C. Storey, "Business intelligence and analytics: From big data to big impact," *MIS Quarterly*, vol. 36, no. 4, pp. 1165–1188, 2012.

[8] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Analysis of recommendation algorithms for E-Commerce," in *Proc. 2nd ACM Conf. Electron. Commerce*, 2000, pp. 158–167.

[9] J. Lampel and H. Mintzberg, "Customizing customization," *MIT Sloan Manag. Rev.*, vol. 38, no. 1, pp. 21–30, 1996.

[10] G. Marcus and E. Davis, "Eight (no, nine!) problems with big data," *The New York Times*, Apr. 6, 2014.

[11] D. Boyd and K. Crawford, "Six provocations for big data," *Social Sci. Res. Netw. Electron. J.*, 2011. [Online]. Available: http://dx.doi.org/10.2139/ssrn.1926431

[12] FDA Clinical Trials Guidance Documents, 2014. [Online]. Available: http://www.fda.gov/regulatoryinformation/guidances/ucm122046.htm

[13] R. Kohavi, "Online controlled experiments," in *Proc. 1st Workshop User Engagement Optim.*, 2013, pp. 15–16.

[14] H. O. Stolberg, G. Norman, and I. Trop, "Randomized controlled trials," *Amer. J. Roentgenol.*, vol. 183, no. 6, pp. 1539–1544, 2014.

[15] D. Lazer *et al.*, "Computational social science," *Science*, vol. 323, no. 5915, pp. 721–723, 2009.

[16] D. Centola and M. Macy, "Complex contagions and the weakness of long ties," *Amer. J. Sociol.*, vol. 113, no. 3, pp. 702–734, 2007.

[17] L. Backstrom and J. Kleinberg, "Network bucket testing," in *Proc. 20th Int. Conf. World Wide Web*, 2011, pp. 615–624.

[18] D. J. Watts and P. S. Dodds, "Influentials, networks, public opinion formation," *J. Consum. Res.*, vol. 34, no. 4, pp. 441–458, Dec. 2007.

[19] P. Dodds and D. Watts, "Universal behavior in a generalized model of contagion," *Phys. Rev. Lett.*, vol. 92, no. 21, 2004, DOI: 10.1103/PhysRevLett.92.218701.

[20] R. Bond *et al.*, "A 61-million-person experiment in social influence and political mobilization," *Nature*, vol. 489, no. 7415, pp. 295–298, 2012.

[21] S. Aral and D. Walker, "Identifying influential and susceptible members of social networks," *Science*, vol. 337, no. 6092, pp. 337–341, Jul. 2012.

[22] S. Aral and D. Walker, "Creating social contagion through viral product design: A randomized trial of peer influence in networks," *Manage. Sci.*, vol. 57, no. 9, pp. 1623–1639, Sep. 2011.

[23] D. Godes, "Firm-created word-of-mouth communication: Evidence from a field test," *Marketing Sci.*, vol. 28, no. 4, pp. 721–739, 2009.

[24] J. Goldenberg, S. Han, D. R. Lehmann, and J. W. Hong, "The role of hubs in the adoption process," *J. Marketing*, vol. 73, no. 2, pp. 1–13, 2009.

[25] R. Iyengar, C. Van den Bulte, and T. W. Valente, "Opinion leadership and social contagion in new product diffusion," *Marketing Sci.*, vol. 30, no. 2, pp. 195–212, 2010.

[26] C. Van den Bulte and S. Wuyts, *Social Networks and Marketing*. Cambridge, MA, USA: Marketing Science Institute, 2007.

[27] A. Goyal, *Social Influence and Its Applications*. Vancouver, BC, Canada: Univ. British Columbia Press, 2013.

[28] E. Bakshy and J. Hofman, "Everyone's an influencer: Quantifying influence on twitter," in *Proc. 4th ACM Int. Conf. Web Search Data Mining*, 2011, pp. 65–74.

[29] M. Webster and J. Sell, *Laboratory Experiments in the Social Sciences*. New York, NY, USA: Academic, 2007, p. 576.

[30] A. Falk and J. Heckman, "Lab experiments are a major source of knowledge in the social sciences," *Science*, vol. 326, no. 5952, pp. 535–538, 2009.

[31] S. Levitt and J. List, "Field experiments in economics: The past, the present, the future," *Eur. Econ. Rev.*, vol. 53, no. 1, pp. 1–18, 2009.

[32] M. Kearns, S. Suri, and N. Montfort, "An experimental study of the coloring problem on human subject networks," *Science*, vol. 313, no. 5788, pp. 824–827, Aug. 2006.

[33] S. Judd, M. Kearns, and Y. Vorobeychik, "Behavioral dynamics and influence in networked coloring and consensus," *Proc. Nat. Acad. Sci. USA*, vol. 107, no. 34, pp. 14978–82, Aug. 2010.

[34] M. Kearns, S. Judd, J. Tan, and J. Wortman, "Behavioral experiments on biased voting in networks," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 5, pp. 1347–1352, Feb. 2009.

[35] L. Qiu, H. Rui, and A. B. Whinston, "The impact of social network structures on prediction market accuracy in the presence of insider information," *J. Manage. Inf. Syst.*, vol. 31, no. 3, pp. 145–172, 2014.

[36] L. Qiu, H. Rui, and A. Whinston, "Effects of social networks on prediction markets: Examination in a controlled experiment," *J. Manage. Inf. Syst.*, vol. 30, no. 4, pp. 235–268, 2014.

[37] G. Paolacci, J. Chandler, and P. Ipeirotis, "Running experiments on Amazon Mechanical Turk," *Judgment Decision Making*, vol. 5, no. 5, pp. 411–419, 2010.

[38] M. J. Salganik, P. S. Dodds, and D. J. Watts, "Experimental study of inequality and unpredictability in an artificial cultural market," *Science*, vol. 311, no. 5762, pp. 854–856, 2006.

[39] S. Suri and D. J. Watts, "Cooperation and contagion in web-based, networked public goods experiments," *PLoS One*, vol. 6, no. 3, Mar. 2011, e16836+.

[40] W. Mason and S. Suri, "Conducting behavioral research on Amazon's Mechanical Turk," *Behav. Res. Methods*, vol. 44, no. 1, pp. 1–23, 2012.

[41] D. Centola, "The spread of behavior in an online social network experiment," *Science*, vol. 329, no. 5996, pp. 1194–1197, Sep. 2010.

[42] D. Centola, "An experimental study of homophily in the adoption of health behavior," *Science*, vol. 334, no. 6060, pp. 1269–1272, Dec. 2011.

[43] D. Centola and A. van de Rijt, "Choosing your network: Social preferences in an online health community," *Social Sci. Med.*, May 2014, DOI: 10.1016/j.socscimed.2014.05.019.

[44] R. Bapna, A. Gupta, S. Rice, and A. Sundararajan, "Trust, reciprocity and the strength of social ties: An online social network based field experiment," working paper, Jul. 2011.

[45] J. Wang, S. Suri, and D. J. Watts, "Cooperation and assortativity with dynamic partner updating," *Proc. Nat. Acad. Sci. USA*, vol. 109, no. 36, pp. 14363–14368, Sep. 2012.

[46] G. Harrison and J. List, "Field experiments," *J. Econ. Lit.*, vol. 42, no. 4, pp. 1009–1055, 2004.

[47] D. Edwards, *I'm Feeling Lucky: The Confessions of Google Employee Number 59*, 1st ed. Chicago, IL, USA: Houghton Mifflin Harcourt, 2011, p. 432.

[48] R. Kohavi *et al.*, "Trustworthy online controlled experiments," in *Proc. 18th ACM SIGKDD Int. Conf. Knowl. Disc. Data Mining*, 2012, pp. 786–794.

[49] E. Bakshy, D. Eckles, and M. Bernstein, "Designing and deploying online field

experiments," in *Proc. 23rd Int. Conf. World Wide Web*, 2014, pp. 283–292.

[50] A. D. I. Kramer, J. E. Guillory, and J. T. Hancock, "Experimental evidence of massive-scale emotional contagion through social networks," *Proc. Nat. Acad. Sci. USA*, vol. 111, no. 24, pp. 8788–9790, Jun. 2014.

[51] S. Harriman and J. Patel, "The ethics and editorial challenges of internet-based research," *BMC Med.*, vol. 12, no. 1, Jul. 2014, DOI:10.1186/s12916-014-0124-3.

[52] R. Bapna, J. Ramaprasad, G. Shmueli, and A. Umyarov, "One-way mirrors and weak-signaling in online dating: A randomized field experiment," working paper, 2013.

[53] E. Bakshy, D. Eckles, R. Yan, and I. Rosenn, "Social influence in social advertising: Evidence from field experiments," in *Proc. 13th ACM Conf. Electron. Commerce*, 2012, pp. 146–161.

[54] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic, "The role of social networks in information diffusion," in *Proc. 21st Int. Conf. World Wide Web*, 2012, pp. 519–528.

[55] S. Taylor, E. Bakshy, and S. Aral, "Selection effects in online sharing: Consequences for peer adoption," in *Proc. 14th ACM Conf. Electron. Commerce*, 2013, pp. 821–836.

[56] S. Taylor, L. Muchnik, and S. Aral, "Identity and opinion: A randomized experiment," working paper, 2014.

[57] S. Aral, L. Muchnik, and A. Sundararajan, "Engineering social contagions: Optimal network seeding and incentive strategies," *Netw. Sci.*, vol. 1, no. 2, pp. 125–153, Feb. 2013.

[58] S. E. Asch, "Effects of group pressure upon the modification and distortion of judgment *Groups, Leadership and Men*. Pittsburgh, PA, USA: Carnegie Press, 1951, pp. 177–190.

[59] S. Milgram, L. Bickman, and L. Berkowitz, "Note on the drawing power of crowds of different size," *J. Personality Social Psychol.*, vol. 13, no. 2, pp. 79–82, 1969.

[60] M. Salganik and D. Watts, "Leading the herd astray: An experimental study of self-fulfilling prophecies in an artificial cultural market," *Social Psychol. Quarterly*, vol. 71, no. 4, pp. 338–355, 2008.

[61] L. Muchnik, S. Aral, and S. Taylor, "Social influence bias: A randomized experiment," *Science*, vol. 341, no. 6146, pp. 647–651, 2013.

[62] M. Godinho de Matos, P. Ferreira, M. D. Smith, and R. Telang, "Culling the herding: Using real world randomized experiments to measure social bias with known costly goods," Social Science Research Network (SSRN) working paper, May 2014.

[63] R. Bapna and A. Umyarov, "Are paid subscriptions on music social networks contagious? A randomized field experiment," Carlson School Manage., Univ. Minnesota, Minneapolis, MN, USA, SOBACO working paper, 2012.

[64] S. Aral and D. Walker, "Tie strength, embeddedness, social influence: A large-scale networked experiment," *Manage. Sci.*, vol. 60, no. 6, pp. 1352–1370, Jun. 2014.

[65] C. F. Manski, "Identification of endogenous social effects: The reflection problem," *Rev. Econ. Studies*, vol. 60, no. 3, pp. 531–542, 1993.

[66] S. Aral, L. Muchnik, and A. Sundararajan, "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks," *Proc. Nat. Acad. Sci USA*, vol. 106, no. 51, pp. 21544–21549, 2009.

[67] C. R. Shalizi and A. C. Thomas, "Homophily and contagion are generically confounded in observational social network studies," *Sociol. Methods Res.*, vol. 42, no. 2, pp. 211–239, 2011.

[68] D. R. Cox, *Planning of Experiments*. New York, NY, USA: Wiley-Interscience, 1958, p. 320.

[69] D. Rubin, "Bayesian inference for causal effects: The role of randomization," *Ann. Stat.*, vol. 6, no. 1, pp. 34–58, 1978.

[70] C. F. Manski, "Identification of treatment response with social interactions," *Econom. J.*, vol. 16, no. 1, pp. S1–S23, Feb. 2013.

[71] P. M. Aronow and C. Samii, "Estimating average causal effects under general interference," in *Proc. Summer Meeting Soc. Political Methodol.*, 2012, pp. 19–21.

[72] D. Eckles, B. Karrer, and J. Ugander, "Design and analysis of experiments in networks: Reducing bias from interference," working paper, 2014, p. 29.

[73] A. Coppock and N. Sircar, "Design of field experiments under unknown interference structures," working paper, 2013.

[74] L. Katzir, E. Liberty, and O. Somekh, "Framework and algorithms for network bucket testing," in *Proc. 21st Int. Conf. World Wide Web*, 2012, pp. 1029–1036.

[75] J. Ugander, B. Karrer, L. Backstrom, and J. Kleinberg, "Graph cluster randomization: Network exposure to multiple universes," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Disc. Data Mining*, 2013, pp. 329–337.

[76] E. Airoldi, P. Toulis, E. Kao, and D. B. Rubin, "Estimation of causal peer influence effects," *J. Mach. Learn. Res.*, vol. 28, no. 3, pp. 1489–1497, 2013.

[77] A. C. Thomas and M. Finegold, "Protocols for randomized experiments to identify network contagion," working paper, 2013, p. 13.

[78] J. Nishimura and J. Ugander, "Restreaming graph partitioning: Simple versatile algorithms for advanced balancing," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Disc. Data Mining*, 2013, pp. 1106–1114.

[79] D. Eckles and E. Bakshy, "Bias and high-dimensional adjustment in observational studies of peer effects," working paper, 2014.

[80] M. J. Salganik and D. D. Heckathorn, "Sampling and estimation in hidden populations using respondent-driven sampling," *Sociol. Methodol.*, vol. 34, no. 1, pp. 193–240, Dec. 2004.

[81] M. Gjoka, M. Kurant, C. T. Butts, and A. Markopoulou, "A walk in Facebook: Uniform sampling of users in online social networks," working paper, May 2009.

[82] D. Godes *et al.*, "The firm's management of social interactions," *Marketing Lett.*, vol. 16, no. 3/4, pp. 415–428, 2005.

[83] S. Aral and S. Taylor, "Viral incentive systems: A randomized field experiment," Stern School Business, New York Univ., working paper, 2014.

[84] G. Pickard *et al.*, "Time-critical social mobilization," *Science*, vol. 334, no. 6055, pp. 509–512, Oct. 2011.

[85] E. Bakshy and D. Eckles, "Uncertainty in online experiments with dependent data: An evaluation of bootstrap methods," in *Proc. 19th ACM SIGKDD Int. Conf. Knowl. Disc. Data Mining*, 2013, pp. 1303–1311.

ABOUT THE AUTHORS

Dylan Walker, received the Ph.D. degree in physics from Stony Brook University, Stony Brook, NY, USA.

He is an Assistant Professor at the School of Management in the Department of Information Systems, Boston University, Boston, MA, USA and a Junior Fellow of the Rafik B. Hariri Institute for Computing and Computational Science & Engineering. Between 2008 and 2012, he was a Research Scientist at the Information, Operations Management and Statistics Department, Leonard N. Stern School of Business, New York University, New York, NY, USA. His research broadly seeks to understand the role of networks and networked systems in the diffusion of information, behaviors, and dynamic processes. His recent work has focused on understanding the dynamics of social influence and peer effects in social networks through large-scale randomized controlled trials. Jointly with collaborators, he has conducted several large-scale randomized controlled trials to understand the dynamics of viral product design and to identify the dependence of peer influence on individual and relationship characteristics in online social networks.

Lev Muchnik, received the Ph.D. degree in physics from Bar Ilan University, Tel Aviv, Israel.

He is a Senior Lecturer at the School of Business Administration, Internet Studies Department Hebrew University of Jerusalem, Jerusalem, Israel. Between 2008 and 2012, he was a Senior Research Scientist at the Information, Operations Management and Statistics Department, Leonard N. Stern School of Business, New York University, New York, NY, USA. His expertise lies in the collection and analysis of massive data sets representing large-scale social systems, and their modeling using tools borrowed from social sciences and statistical physics. His recent work has been focused on theoretical and empirical problems related to the structure and evolution of social networks, as well as peer effects, the spread of behavioral norms, information diffusion, and other processes specific to networked environments. Jointly with collaborators, he developed a seminal method for the identification of peer influence in networks, and conducted large-scale randomized controlled experiments in online communities.

FUTURE ISSUES

The following is a schedule of upcoming special issues and sections.

2015
Global Health Care: Advances and Challenges
Guest Editors:
Metin Akay
Department of Biomedical Engineering
University of Houston
Houston, TX, USA

2015
Design Automation of Electronic Systems—Past Accomplishments and Challenges Ahead
Guest Editors:
Robert Brayton
Department of Electrical Engineering and Computer Science
University of California
Berkeley, CA, USA

Luca Carloni
Department of Computer Science
Columbia University
New York, NY, USA

Alberto Sangiovanni-Vincentelli
Department of Electrical Engineering and Computer Science
University of California
Berkeley, CA, USA

Tiziano Villa
Dipartimento d'Informatica
Universita' di Verona
Verona, Italy

2015
Reconfigurable Systems
Guest Editors:
Christos Christodoulou
COSMIAC/University of New Mexico
Albuquerque, NM, USA

James Lyke
Air Force Research Laboratory
Kirtland Air Force Base, NM, USA

Alonzo Vera
University of New Mexico
Albuquerque, NM, USA

Arthur Edwards
Air Force Research Laboratory
Albuquerque, NM, USA

2015
The Plurality of Human Brain Computer Interfaces
Guest Editors:
Gernot Müller-Putz
Graz University of Technology
Graz, Austria

José del R. Millán
Ecole Polytechnique Fédérale de Lausanne (EPFL)
Lausanne, Switzerland

Gerwin Schalk
Wadsworth Center
NYS Department of Health
Albany, NY, USA

Klaus-Robert Müller
Machine Learning Group
Technical University Berlin
Berlin, Germany

2015
Physical-Layer Security and its Applications
Guest Editors:
Phillip A. Regalia
National Science Foundation
USA

Yingbin Liang
Syracuse University
Syracuse, NY, USA

Ashish Khisti
University of Toronto
Toronto, ON, Canada

Stefano Tomasin
University of Padova
Padova, Italy

2015
Multimodal Data Fusion
Guest Editors:
Tülay Adali
North Carolina State University
Raleigh, NC, USA

Christian Jutten
Institut Universitaire de France
University of Grenoble
Grenoble, France

Lars Kai Hansen
Technical University of Denmark
Kongens Lyngby, Denmark

2015
Advanced Flexible Electronics for Sensing Applications
Guest Editors:
Bruce Gnade
University of Texas at Dallas
Dallas, TX, USA

Gregory B. Raupp
MacroTechnology Works Initiative
Arizona State University
Tempe, AZ, USA

Robert H. Reuss
Independent Consultant
Flexible/Large Area Electronics Organizations

2015
Memories in the Future of Information Processing
Guest Editor:
Sandip Tiwari
Cornell University
Ithaca, NY, USA

2015
Big Data: Science Perspective
Guest Editors:
Simon Haykin
MacMaster University
Hamilton, ON, Canada

Yoshua Benigo
University of Montreal
Montreal, QC, Canada

Stephen Wright
University of Wisconsin
Madison, WI, USA

Jon Atli Benediktsson
University of Iceland
Reykjavik, Iceland

2015
Microgrids and Energy Sustainable Buildings
Guest Editors:
John Baillieul
Boston University
Boston, MA, USA

Michael Caramanis
Boston University
Boston, MA, USA

Marija Ilic
Carnegie Mellon University
Pittsburgh, PA, USA

2015
Electric Ship Technologies
Guest Editors:
Steinar J. Dale
Florida State University
Tallahassee, FL, USA

Robert E. Hebner
University of Texas at Austin
Austin, TX, USA

Giorgio Sulligoi
University of Trieste
Trieste, Italy

2015
Structural Health Monitoring: Advances and Implementations
Guest Editors:
Jerome P. Lynch
University of Michigan
Ann Arbor, MI, USA

Jennifer E. Michaels
Georgia Institute of Technology
Atlanta, GA, USA

Charles R. Farrar
Los Alamos National Laboratory
Los Alamos, NM, USA

2015
Informatics for Nanoscience: Developing New Materials for the Future
Guest Editors:
Krishna Rajan
Director
Institute for Combinatorial Discovery
Iowa State University
Ames, IA, USA

Sadasivan Shankar
Senior Principal Engineer and Program Leader for Materials Design
Design and Technology Group
Intel Technology and Manufacturing Organization
San Jose, CA, USA