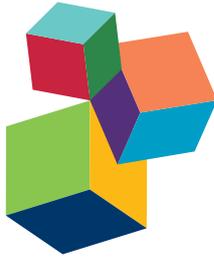


AT THE CROSSROADS: LESSONS AND CHALLENGES IN COMPUTATIONAL SOCIAL SCIENCE

EDITED BY: Javier Borge-Holthoefer, Yamir Moreno and Taha Yasseri
PUBLISHED IN: Frontiers in Physics



frontiers

Frontiers Copyright Statement

© Copyright 2007-2016 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-021-3

DOI 10.3389/978-2-88945-021-3

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.

Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.

By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

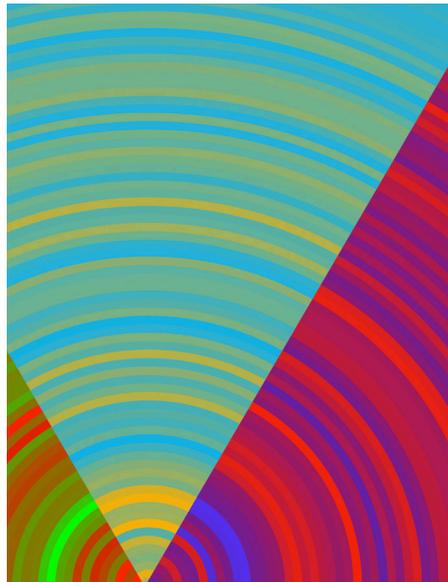
AT THE CROSSROADS: LESSONS AND CHALLENGES IN COMPUTATIONAL SOCIAL SCIENCE

Topic Editors:

Javier Borge-Holthoefer, Universitat Oberta de Catalunya, Spain & Hamad Bin Khalifa University, Qatar & Universidad de Zaragoza, Spain

Yamir Moreno, Universidad de Zaragoza, Spain & Institute for Scientific Interchange, Italy

Taha Yasseri, University of Oxford & Alan Turing Institute, UK



Cover Image shows emotional expression in the trust network of a product reviews community. The users in a central core of the network are more emotional than the rest. Read the chapter by Tanase et al. (<http://dx.doi.org/10.3389/fphy.2015.00087>).

The interest of physicists in economic and social questions is not new: for over four decades, we have witnessed the emergence of what is called nowadays “sociophysics” and “econophysics”, vigorous and challenging areas within the wider “Interdisciplinary Physics”. With tools borrowed from Statistical Physics and Complexity, this new area of study has already made important contributions, which in turn have fostered the development of novel theoretical foundations in Social Science and Economics, via mathematical approaches, agent-based modelling and numerical simulations.

From these foundations, Computational Social Science has grown to incorporate as well the empirical component —aided by the recent data deluge from the Web 2.0 and 3.0—, closing in this way the experiment-theory cycle in the best tradition of Physics.

Citation: Borge-Holthoefer, J., Moreno, Y., Yasseri, T., eds. (2016). *At the Crossroads: Lessons and Challenges in Computational Social Science*. Lausanne: Frontiers Media.
doi: 10.3389/978-2-88945-021-3

Table of Contents

- 04 Editorial: At the Crossroads: Lessons and Challenges in Computational Social Science**
Javier Borge-Holthoefer, Yamir Moreno and Taha Yasseri
- 07 P-Values: Misunderstood and Misused**
Bertie Vidgen and Taha Yasseri
- 12 Mechanistic models in computational social science**
Petter Holme and Fredrik Liljeros
- 22 Citizen Science Practices for Computational Social Science Research: The Conceptualization of Pop-Up Experiments**
Oleguer Sagarra, Mario Gutiérrez-Roig, Isabelle Bonhoure and Josep Perelló
- 36 Mathematical modeling of complex contagion on clustered networks**
David J. P. O'Sullivan, Gary J. O'Keeffe, Peter G. Fennell and James P. Gleeson
- 49 Information transfer in community structured multiplex networks**
Albert Solé-Ribalta, Clara Granell, Sergio Gómez and Alex Arenas
- 56 Temporal pattern of online communication spike trains in spreading a scientific rumor: how often, who interacts with whom?**
Ceyda Sanli and Renaud Lambiotte
- 64 Characterizing interactions in online social networks during exceptional events**
Elisa Omodei, Manlio De Domenico and Alex Arenas
- 72 Digital daily cycles of individuals**
Talayah Aledavood, Sune Lehmann and Jari Saramäki
- 79 Emotions and Activity Profiles of Influential Users in Product Reviews Communities**
Dorian Tanase, David Garcia, Antonios Garas and Frank Schweitzer
- 91 A Biased Review of Biases in Twitter Studies on Political Collective Action**
Peter Cihon and Taha Yasseri



Editorial: At the Crossroads: Lessons and Challenges in Computational Social Science

Javier Borge-Holthoefer^{1,2,3*}, Yamir Moreno^{3,4,5} and Taha Yasseri^{6,7}

¹ Complex Systems Group (CoSIN3), Internet Interdisciplinary Institute (IN3), Universitat Oberta de Catalunya, Barcelona, Spain, ² Qatar Computing Research Institute, Hamad Bin Khalifa University, Doha, Qatar, ³ Institute of Biocomputation and Physics of Complex Systems, Universidad de Zaragoza, Zaragoza, Spain, ⁴ Department of Theoretical Physics, Faculty of Sciences, Universidad de Zaragoza, Zaragoza, Spain, ⁵ Institute for Scientific Interchange, Torino, Italy, ⁶ Oxford Internet Institute, University of Oxford, Oxford, UK, ⁷ Alan Turing Institute, London, UK

Keywords: computational social science, simulation, models, big data, complex systems

The Editorial on the Research Topic

At the Crossroads: Lessons and Challenges in Computational Social Science

The interest of physicists in economic and social questions is not new: during the last decades, we have witnessed the emergence of what is formally called nowadays sociophysics [1] and econophysics [2] that can be grouped into the common term “Interdisciplinary Physics” along with biophysics, medical physics, agrophysics, etc. With tools borrowed from statistical physics and complexity science, among others, these areas of study have already made important contributions to our understanding of how humans organize and interact in our modern society. Large scale data analyses, agent-based modeling and numerical simulations, and finally mathematical modeling, have led to the discovery of new (universal) patterns and their quantitative description in socio-economic systems.

At the turn of the century, however, it was clear that huge challenges—and new opportunities—lied ahead: the digital communication technologies, and their associated data deluge, began to nurture those models with empirical significance. Only a decade later, the advent of the Web 2.0, the Internet of Things and a general adoption of mobile technologies have convinced researchers that theories can be mapped to real scenarios and put into empirical test, closing in this way the experiment-theory cycle in the best tradition of physics.

We are nowadays at a crossroads, at which different approaches converge. We name such crossroads computational social science (CSS) : a new discipline that can offer abstracted (simplified, idealized) models and methods (mainly from statistical physics), large storage, algorithms and computational power (computer and data science), and a set of social hypotheses together with a conceptual framework for the results to be interpreted (Social Science) [3–5]. Despite its youth, the field is developing rapidly in terms of contents (articles, books, etc.), but also institutionally—either under the form of labs, institutes, and academic programs; or as consolidated events and scientific gatherings.

This “work-in-progress” spirit is reflected as well in this volume: the call was launched in late 2014 and 10 articles were eventually accepted and published, including reviews—a look behind—, one methods paper, and six original contributions—a look ahead—introducing a broad range of research, from models with a strong analytical flavor to data-driven problems.

As mentioned above, each new research line in CSS starts with analysing a sizable dataset, containing transactional data or user generated content on social web (also called Big Data), The availability of data however poses new methodological challenges. Among them are statistical

OPEN ACCESS

Edited and reviewed by:

Alex Hansen,
Norwegian University of Science and
Technology, Norway

*Correspondence:

Javier Borge-Holthoefer
jborgeh@uoc.edu

Specialty section:

This article was submitted to
Interdisciplinary Physics,
a section of the journal
Frontiers in Physics

Received: 29 July 2016

Accepted: 16 August 2016

Published: 29 August 2016

Citation:

Borge-Holthoefer J, Moreno Y and
Yasseri T (2016) Editorial: At the
Crossroads: Lessons and Challenges
in Computational Social Science.
Front. Phys. 4:37.
doi: 10.3389/fphy.2016.00037

analysis and how the methods and conventions that have been used in social sciences to analyse small datasets collected via questionnaires and interviews, can be now used to analyse data generated by millions of people. Vidgen and Yasserli address this challenge and in particular discuss the confusion and (mis)uses around the highly-popular p -values. They call for a more careful use of statistical tests and show few directions for improvement. Apart from the methodological challenges, as Holme and Liljeros note in their review on mechanistic models in computational social science, “Quantitative social science is not only about regression analysis or, in general, data inference.” Computer simulations, whose history is reviewed by Holme and Liljeros, are one the main connecting bridges between empirical observations and theoretical models.

Adopted from physics, the next step in the “scientific method” of CSS is experimentation. Experiments might not only be used to validate the generated theories, but also to yield new observations that might eventually lead to new lines of research themselves. However, experimentation in social content and with human subjects is nothing similar to physics experiments. Recruitment, representativeness, privacy issues, and ethical challenges are very central to social experimentation. Sagarra et al., focus their attention on citizen science and offer a methodological guideline for experimentation outside the laboratory.

Spreading phenomena make one of the key topics in CSS. With application in innovation [6], political change [7], epidemiology [8], etc., it is important to understand how “things” navigate through social networks. To further develop the extensive literature of social contagion, O’Sullivan et al. extend the modeling of complex contagion in the context of clustered networks—of upmost relevance in the social context. Solé-Ribalta et al., on their side, focus on communities on multilayer networks [9] and develop new mathematics for information transfer. These two articles make an excellent case of how the insights from complex systems and statistical physics can and do play a role in CSS, offering solid foundations to the methods and insights developed within.

One of the main advantages of CSS over classic social sciences, is the possibility to perform temporal analysis and come up with dynamical models. Most of the datasets under study in CSS contain timestamps, allowing for fine-grained analysis of interactions over time. Sanli and Lambiotte provide an original approach to online communication based on complex time series—rather than on network structure— that emerge from user dynamics on social media. The analysis is performed on a set

of collected messages which correspond to an exceptional event, which is common practice in the field to study collective behavior [10]. Also Omodei et al. take this approach, analyzing a broad range of events (policy, culture, science) which they characterize as multiplex networks [9, 11]. And finally Aledavood et al. use mobile phone records to study diurnal patterns of human communication and provide a cohesive picture of regularities in communication patterns both at individual and society levels.

Large scale analysis of socially generated data is not limited to transactional records: huge amount of digital content is being produced on daily basis. In a novel work, Tanase et al. apply linguistic analysis to user reviews that they collected from the web and study social influence and its interplay with network topology, and how it affects users’ opinions.

Generalizing on opinion dynamics, there is no doubt that the Internet in general and social media in particular have changed the political environment and the way people engage in political activities [12]. At the same time, the digital footprint of online political activities provides a great opportunity to conduct political science studies at scale and in close-to real time, leading to the emergence of some sort of data-driven political science within CSS [13]. However, these opportunities come with challenges and shortcomings. Cihon and Yasserli take a critical point of view toward such studies and in particular discuss the “biases” in Twitter-based research on political collective action in a short “biased review.”

Computational Social Science emerges as a wide set of scientific opportunities, to tackle the fundamental features of social complexity—multidirectional connections, layer interdependences and interferences, accelerated diffusion, and so on [14]. The complex systems approach that underlies CSS is a key feature toward creating a truly interdisciplinary, non-compartmental science.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

We thank Matjaž Perc for the great editorial help and all the reviewers who assisted us in preparing the manuscripts of this Research Topic.

REFERENCES

- Sen P, Chakrabarti BK. *Sociophysics: An Introduction*. Oxford, UK: Oxford University Press (2013).
- Mantegna RN, Stanley HE. *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge, UK: Cambridge University Press (1999). doi: 10.1017/CBO9780511755767
- Lazer D, Pentland AS, Adamic L, Aral S, Barabasi AL, Brewer D, et al. Life in the network: the coming age of computational social science. *Science* (2009) 323:721. doi: 10.1126/science.1167742
- Watts DJ. Computational social science: exciting progress and future directions. *Bridge Front. Eng.* (2013) 43:5–10. Available online at: <https://www.nae.edu/Publications/Bridge/106112/106118.aspx>
- Mann A. Core concepts: computational social science. *Proc Natl Acad Sci USA* (2016) 113:468–70. doi: 10.1073/pnas.1524881113
- Pentland A. *Social Physics: How Good Ideas Spread—the Lessons from a New Science*. New York, NY: Penguin (2014).
- Galam S. *Sociophysics: A Physicist’s Modeling of Psycho-Political Phenomena*. New York, NY: Springer Science & Business Media (2012).
- Newman ME. Spread of epidemic disease on networks. *Phys Rev E* (2002) 66:016128. doi: 10.1103/PhysRevE.66.016128

9. Kivela M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA. Multilayer networks. *J Comp Netw.* (2014) 2:203–71. doi: 10.1093/comnet/cnu016
10. Borge-Holthoefer J, Perra N, Gonçalves B, González-Bailón S, Arenas A, Moreno Y, et al. The dynamics of information-driven coordination phenomena: a transfer entropy analysis. *Sci Adv.* (2016) 2:e1501158. doi: 10.1126/sciadv.1501158
11. De Domenico M, Solé-Ribalta A, Cozzo E, Kivela M, Moreno Y, Porter MA, et al. Mathematical formulation of multilayer networks. *Phys Rev X* (2013) 3:041022. doi: 10.1103/PhysRevX.3.041022
12. Margetts H, John P, Hale S, Yasseri T. *Political Turbulence: How Social Media Shape Collective Action*. Princeton, NJ: Princeton University Press (2015).
13. Weber I, Popescu AM, Pennacchiotti M. Data-driven political science. In: *Proceedings of the Sixth ACM International Conference on Web Search and Data Mining*. New York, NY: ACM (2013). p. 777–778. doi: 10.1145/2433396.2433498
14. Conte R, Gilbert N, Bonelli G, Cioffi-Revilla C, Deffuant G, Kertesz J, et al. Manifesto of computational social science. *Eur Phys J Spec Top.* (2012) 214:325–46. doi: 10.1140/epjst/e2012-01697-8

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Borge-Holthoefer, Moreno and Yasseri. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



P-Values: Misunderstood and Misused

*Bertie Vidgen and Taha Yasseri**

Oxford Internet Institute, University of Oxford, Oxford, UK

P-values are widely used in both the social and natural sciences to quantify the statistical significance of observed results. The recent surge of big data research has made the *p*-value an even more popular tool to test the significance of a study. However, substantial literature has been produced critiquing how *p*-values are used and understood. In this paper we review this recent critical literature, much of which is routed in the life sciences, and consider its implications for social scientific research. We provide a coherent picture of what the main criticisms are, and draw together and disambiguate common themes. In particular, we explain how the False Discovery Rate (FDR) is calculated, and how this differs from a *p*-value. We also make explicit the Bayesian nature of many recent criticisms, a dimension that is often underplayed or ignored. We conclude by identifying practical steps to help remediate some of the concerns identified. We recommend that (i) far lower significance levels are used, such as 0.01 or 0.001, and (ii) *p*-values are interpreted contextually, and situated within both the findings of the individual study and the broader field of inquiry (through, for example, meta-analyses).

OPEN ACCESS

Edited by:

Matjaž Perc,
University of Maribor, Slovenia

Reviewed by:

Haroldo Valentin Ribeiro,
Universidade Estadual de Maringá,
Brazil
Megan Head,
Australian National University, Australia

*Correspondence:

Taha Yasseri
taha.yasseri@oii.ox.ac.uk

Specialty section:

This article was submitted to
Interdisciplinary Physics,
a section of the journal
Frontiers in Physics

Received: 25 January 2016

Accepted: 19 February 2016

Published: 04 March 2016

Citation:

Vidgen B and Yasseri T (2016)
*P-Values: Misunderstood and
Misused. Front. Phys. 4:6.*
doi: 10.3389/fphy.2016.00006

Keywords: *p*-value, statistics, significance, p-hacking, prevalence, Bayes, big data

1. INTRODUCTION

P-values are widely used in both the social and natural sciences to quantify the statistical significance of observed results. Obtaining a *p*-value that indicates “statistical significance” is often a requirement for publishing in a top journal. The emergence of computational social science, which relies mostly on analyzing large scale datasets, has increased the popularity of *p*-values even further. However, critics contend that *p*-values are routinely misunderstood and misused by many practitioners, and that even when understood correctly they are an ineffective metric: the standard significance level of 0.05 produces an overall FDR that is far higher, more like 30%. Others argue that *p*-values can be easily “hacked” to indicate statistical significance when none exists, and that they encourage the selective reporting of only positive results.

Considerable research exists into how *p*-values are (mis)used, [e.g., 1, 2]. In this paper we review the recent critical literature on *p*-values, much of which is routed in the life sciences, and consider its implications for social scientific research. We provide a coherent picture of what the main criticisms are, and draw together and disambiguate common themes. In particular, we explain how the FDR is calculated, and how this differs from a *p*-value. We also make explicit the Bayesian nature of many recent criticisms. In the final section we identify practical steps to help remediate some of the concerns identified.

P-values are used in Null Hypothesis Significance Testing (NHST) to decide whether to accept or reject a null hypothesis (which typically states that there is no underlying relationship between two variables). If the null hypothesis is rejected, this gives grounds for accepting the alternative

hypothesis (that a relationship does exist between two variables). The p -value quantifies the probability of observing results at least as extreme as the ones observed given that the null hypothesis is true. It is then compared against a pre-determined significance level (α). If the reported p -value is smaller than α the result is considered statistically significant. Typically, in the social sciences α is set at 0.05. Other commonly used significance levels are 0.01 and 0.001.

In his seminal paper, “The Earth is Round ($p < .05$)” Cohen argues that NHST is highly flawed: it is relatively easy to achieve results that can be labeled significant when a “nil” hypothesis (where the effect size of H_0 is set at zero) is used rather than a true “null” hypothesis (where the direction of the effect, or even the effect size, is specified) [3]. This problem is particularly acute in the context of “big data” exploratory studies, where researchers only seek statistical associations rather than causal relationships. If a large enough number of variables are examined, effectively meaning that a large number of null/alternative hypotheses are specified, then it is highly likely that at least some “statistically significant” results will be identified, irrespective of whether the underlying relationships are truly meaningful. As big data approaches become more common this issue will become both far more pertinent and problematic, with the robustness of many “statistically significant” findings being highly limited.

Low argues that the central problem with NHST is reflected in its hybrid name, which is a combination of (i) hypothesis testing and (ii) significance testing [4]. In significance testing, first developed by Ronald Fisher in the 1920s, the p -value provides an index of the evidence against the null hypothesis. Originally, Fisher only intended for the p -value to establish whether further research into a phenomenon could be justified. He saw it as one bit of evidence to either support or challenge accepting the null hypothesis, rather than as conclusive evidence of significance [5; see also 6, 7]. In contrast, hypothesis tests, developed separately by Neyman and Pearson, replace Fisher’s subjectivist interpretation of the p -value with a hard and fast “decision rule”: when the p -value is less than α , the null can be rejected and the alternative hypothesis accepted. Though this approach is simpler to apply and understand, a crucial stipulation of it is that a precise alternative hypothesis must be specified [6]. This means indicating what the expected effect size is (thereby setting a nil rather than a null hypothesis)—something that most researchers rarely do [3].

Though hypothesis tests and significance tests are distinct statistical procedures, and there is much disagreement about whether they can be reconciled into one coherent framework, NHST is widely used as a pragmatic amalgam for conducting research [8, 9]. Hulbert and Lombardi argue that one of the biggest issues with NHST is that it encourages the use of terminology such as *significant/nonsignificant*. This dichotomizes the p -value on an arbitrary basis, and converts a probability into a certainty. This is unhelpful when the purpose of using statistics, as is typically the case in academic studies, is to weigh up evidence incrementally rather than make an immediate decision [9, p. 315]. Hulbert and Lombardi’s analysis suggests that the real problem lies not with p -values, but with α and how this has led to

p -values being interpreted dichotomously: too much importance is attached to the arbitrary cutoff $\alpha \leq 0.05$.

2. THE FALSE DISCOVERY RATE

A p -value of 0.05 is normally interpreted to mean that there is a 1 in 20 chance that the observed results are nonsignificant, having occurred even though no underlying relationship exists. Most people then think that the overall proportion of results that are false positives is also 0.05. However, this interpretation confuses the p -value (which, in the long run, will approximately correspond to the *type I error rate*) with the FDR. The FDR is what people usually mean when they refer to the error rate: it is the proportion of reported discoveries that are false positives. Though 0.05 might seem a reasonable level of inaccuracy, a type I error rate of 0.05 will likely produce an FDR that is far higher, easily 30% or more. The formula for FDR is:

$$\frac{\text{False Positives}}{\text{True Positives} + \text{False Positives}} \quad (1)$$

Calculating the number of true positives and false positives requires knowing more than just the type I error rate, but also (i) the statistical power, or “sensitivity,” of tests and (ii) the prevalence of effects [10]. Statistical power is the probability that each test will correctly reject the null hypothesis when the alternative hypothesis is true. As such, tests with higher power are more likely to correctly record real effects. Prevalence is the number of effects, out of all the effects that are tested for, that actually exist in the real world. In the FDR calculation it determines the weighting given to the power and the type I error rate. Low prevalence contributes to a higher FDR as it increases the likelihood that false positives will be recorded. The calculation for FDR therefore is:

$$\frac{(1 - \text{Prevalence}) \times \text{Type I error rate}}{\text{Prevalence} \times \text{Power} + (1 - \text{Prevalence}) \times \text{Type I error rate}} \quad (2)$$

The percentage of reported positives that are actually true is called the Positive Predictive Value (PPV). The PPV and FDR are inversely related, such that a higher PPV necessarily means a lower FDR. To calculate the FDR we subtract the PPV from 1. If there are no false positives then $\text{PPV} = 1$ and $\text{FDR} = 0$. **Table 1** shows how low prevalence of effects, low power, and a high type I error rate all contribute to a high FDR.

Most estimates of the FDR are surprisingly large; e.g., 50 [1, 11, 12] or 36% [10]. Jager and Leek more optimistically suggest that it is just 14% [13]. This lower estimate can be explained somewhat by the fact that they only use p -values reported in abstracts, and have a different algorithm to the other studies. Importantly, they highlight that whilst α is normally set to 0.05, many studies—particularly in the life sciences—achieve p -values far lower than this, meaning that the average type I error rate is less than α of 0.05 [13, p. 7]. Counterbalancing this, however, is Colquhoun’s argument that because most studies are not “properly designed” (in the sense that treatments are not randomly allocated to groups and in RCTs assessments are not blinded) statistical power will

TABLE 1 | Greater prevalence, greater power, and a lower Type I error rate reduce the FDR.

Prevalence	Power	Type I error rate	FDR
0.01	0.8	0.05	0.86
0.1	0.8	0.05	0.36
0.5	0.8	0.05	0.06
0.1	0.2	0.05	0.69
0.1	0.5	0.05	0.47
0.1	0.8	0.01	0.10
0.1	0.8	0.001	0.01

often be far lower than reported—thereby driving the FDR back up again [10].

Thus, though difficult to calculate precisely, the evidence suggests that the FDR of findings overall is far higher than α of 0.05. This suggests that too much trust is placed in current research, much of which is wrong far more often than we think. It is also worth noting that this analysis assumes that researchers do not intentionally misreport or manipulate results to erroneously achieve statistical significance. These phenomena, known as “selective reporting” and “p-hacking,” are considered separately in Section 4.

3. PREVALENCE AND BAYES

As noted above, the prevalence of effects significantly impacts the FDR, whereby lower prevalence increases the likelihood that reported effects are false positives. Yet prevalence is not controlled by the researcher and, furthermore, cannot be calculated with any reliable accuracy. There is no way of knowing objectively what the underlying prevalence of real effects is. Indeed, the tools by which we might hope to find out this information (such as NHST) are precisely what have been criticized in the literature surveyed here. Instead, to calculate the FDR, prevalence has to be estimated¹. In this regard, FDR calculations are inherently Bayesian as they require the researcher to quantify their subjective belief about a phenomenon (in this instance, the underlying prevalence of real effects).

Bayesian theory is an alternative paradigm of statistical inference to frequentism, of which NHST is part of. Whereas, frequentists quantify the probability of the data given the null hypothesis ($P(D|H_0)$), Bayesians calculate the probability of the hypothesis given the data ($P(H_1|D)$). Though frequentism is far more widely practiced than Bayesianism, Bayesian inference is more intuitive: it assigns a probability to a hypothesis based on how likely we think it to be true.

The FDR calculations outlined above in Section 2 follow a Bayesian logic. First, a probability is assigned to the prior likelihood of a result being false ($1 - \text{prevalence}$). Then, new information (the statistical power and type I error rate) is incorporated to calculate a posterior probability (the FDR). A

¹In much of the recent literature it is assumed that prevalence is very low, around 0.1 or 0.2 [1, 10–12].

common criticism against Bayesian methods such as this is that they are insufficiently objective as the prior probability is only a guess. Whilst this is correct, the large number of “findings” produced each year, as well as the low rates of replicability [14], suggest that the prevalence of effects is, overall, fairly low. Another criticism against Bayesian inference is that it is overly conservative: assigning a low value to the prior probability makes it more likely that the posterior probability will also be low [15]. These criticisms notwithstanding, Bayesian theory offers a useful way of quantifying how likely it is that research findings are true.

Not all of the authors in the literature reviewed here explicitly state that their arguments are Bayesian. The reason for this is best articulated by Colquhoun, who writes that “the description ‘Bayesian’ is not wrong but it is not necessary” [10, p. 5]. The lack of attention paid to Bayes in Ioannidis’ well-regarded early article on *p*-values is particularly surprising given his use of Bayesian terminology: “the probability that a research finding is true depends on the prior probability of it being true (before doing the study)” [1, p. 696]. This perhaps reflects the uncertain position that Bayesianism holds in most universities, and the acrimonious nature of its relationship with frequentism [16]. Without commenting on the broader applicability of Bayesian statistical inference, we argue that a Bayesian methodology has great utility in assessing the overall credibility of academic research, and that it has received insufficient attention in previous studies. Here, we have sought to make visible, and to rectify, this oversight.

4. PUBLICATION BIAS: SELECTIVE REPORTING AND P-HACKING

Selective reporting and p-hacking are two types of researcher-driven publication bias. Selective reporting is where nonsignificant (but methodologically robust) results are not reported, often because top journals consider them to be less interesting or important [17]. This skews the distribution of reported results toward positive findings, and arguably further increases the pressure on researchers to achieve statistical significance. Another form of publication bias, which also skews results toward positive findings, is called p-hacking. Head et al. define p-hacking as “when researchers collect or select data or statistical analyses until nonsignificant results become significant” [18]. This is direct manipulation of results so that, whilst they may not be technically false, they are unrepresentative of the underlying phenomena. See **Figure 1** for a satirical illustration.

Head et al. outline specific mechanisms by which *p*-values are intentionally “hacked.” These include: (i) conducting analyses midway through experiments, (ii) recording many response variables and only deciding which to report postanalysis, (iii) excluding, combining, or splitting treatment groups postanalysis, (iv) including or excluding covariates postanalysis, (v) stopping data exploration if analysis yields a significant *p*-value. An excellent demonstration of how *p*-values can be hacked by manipulating the parameters of an experiment is Christie Aschwanden’s interactive “Hack Your Way to Scientific Glory”

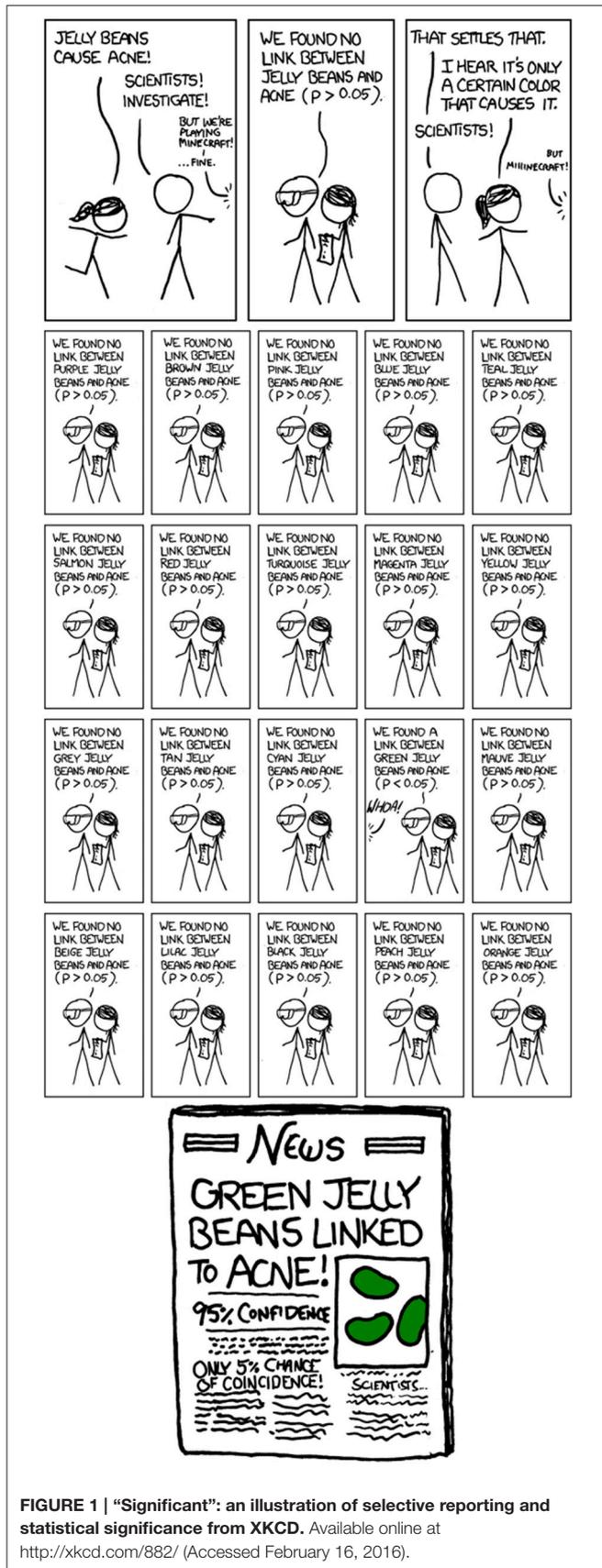


FIGURE 1 | “Significant”: an illustration of selective reporting and statistical significance from XKCD. Available online at <http://xkcd.com/882/> (Accessed February 16, 2016).

[19]. This simulator, which analyses whether Republicans or Democrats being in office affects the US economy, shows how tests can be manipulated to produce statistically significant results supporting either parties.

In separate papers, Head et al. [18], and de Winter and Dodou [20] each examine the distributions of p -values that are reported in scientific publications in different disciplines. It is reported that there are considerably more studies reporting alpha just below the 0.05 significance level than above it (and considerably more than would be expected given the number of p -values that occur in other ranges), which suggests that p -hacking is taking place. This core finding is supported by Jager and Leek’s study on “significant” publications as well [13].

5. WHAT TO DO

We argued above that a Bayesian approach is useful to estimate the FDR and assess the overall trustworthiness of academic findings. However, this does not mean that we also hold that Bayesian statistics should replace frequentist statistics more generally in empirical research [see: 21]. In this concluding section we recommend some pragmatic changes to current (frequentist) research practices that could lower the FDR and thus improve the credibility of findings.

Unfortunately, researchers cannot control how prevalent effects are. They only have direct influence over their study’s α and its statistical power. Thus, one step to reduce the FDR is to make the norms for these more rigorous, such as by increasing the statistical power of studies. We strongly recommend that α of 0.05 is dropped as a convention, and replaced with a far lower α as standard, such as 0.01 or 0.001; see **Table 1**. Other suggestions for improving the quality of statistical significance reporting include using confidence intervals [7, p. 152]. Some have also called for researchers to focus more on effect sizes than statistical significance [22, 23], arguing that statistically significant studies that have negligible effect sizes should be treated with greater skepticism. This is of particular importance in the context of big data studies, where many “statistically significant” studies report small effect sizes as the association between the dependent and independent variables is very weak.

Perhaps more important than any specific technical change in how data is analyzed is the growing consensus that research processes need to be implemented (and recorded) more transparently. Nuzzo, for example, argues that “one of the strongest protections for scientists is to admit everything” [7, p. 152]. Head et al. also suggest that labeling research as either exploratory or confirmatory will help readers to interpret the results more faithfully [18, p. 12]. Weissgerber et al. encourage researchers to provide “a more complete presentation of data,” beyond summary statistics [24]. Improving transparency is particularly important in “big” data-mining studies, given that the boundary between data exploration (a legitimate exercise) and p -hacking is often hard to identify, creating significant potential for intentional or unintentional manipulation of results. Several commentators have recommended that researchers pre-register all studies with initiatives such as the Open Science

Framework [1, 7, 14, 18, 25]. Pre-registering ensures that a record is kept of the proposed method, effect size measurement, and what sort of results will be considered noteworthy. Any deviation from what is initially registered would then need to be justified, which would give the results greater credibility. Journals could also proactively assist researchers to improve transparency by providing platforms on which data and code can be shared, thus allowing external researchers to reproduce a study's findings and trace the method used [18]. This would provide academics with the practical means to corroborate or challenge previous findings.

Scientific knowledge advances through corroboration and incremental progress. In keeping with Fisher's initial view that p -values should be one part of the evidence used when deciding whether to reject the null hypothesis, our final suggestion is that the findings of any single study should always be contextualized within the broader field of research. Thus, we endorse the view offered in a recent editorial of *Psychological Science* that we should be extra skeptical about studies where (a) the statistical power is low, (b) the p -value is only slightly below 0.05, and (c) the result is surprising [14]. Normally, findings are only accepted

once they have been corroborated through multiple studies, and even in individual studies it is common to "triangulate" a result with multiple methods and/or data sets. This offers one way of remediating the problem that even "statistically significant" results can be false; if multiple studies find an effect then it is more likely that it truly exists. We therefore, also support the collation and organization of research findings in meta-analyses as these enable researchers to quickly evaluate a large range of relevant evidence.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

For providing useful feedback on the original manuscript we thank Jonathan Bright, Sandra Wachter, Patricia L. Mabry, and Richard Vidgen.

REFERENCES

- Ioannidis J. Why most published research findings are false. *PLoS Med.* (2005) 2:e124. doi: 10.1371/journal.pmed.0020124
- Ziliak ST, McCloskey DN. *The Cult of Statistical Significance: How the Standard Error Costs us Jobs, Justice, and Lives.* Ann Arbor, MI: University of Michigan Press (2008).
- Cohen J. The earth is round ($p < .05$). *Am Psychol.* (1994) 49:997–1003.
- Lew MJ. To P or not to P: on the evidential nature of P-values and their place in scientific inference. arXiv preprint arXiv:13110081 (2013).
- Fisher RA. *Statistical Methods for Research Workers.* Edinburgh: Genesis Publishing Pvt. Ltd. (1925).
- Sterne JA, Smith GD. Sifting the evidence—what's wrong with significance tests? *BMJ* (2001) 322:226–31. doi: 10.1136/bmj.322.7280.226
- Nuzzo R. Statistical errors. *Nature* (2014) 506:150–2. doi: 10.1038/506150a
- Berger JO. Could fisher, jeffreys and neyman have agreed on testing? *Stat Sci.* (2003) 18:1–32. doi: 10.1214/ss/1056397485
- Hurlbert SH, Lombardi CM. Final collapse of the Neyman-Pearson decision theoretic framework and rise of the neoFisherian. *Ann Zool Fenn.* (2009) 46:311–49. doi: 10.5735/086.046.0501
- Colquhoun D. An investigation of the false discovery rate and the misinterpretation of p -values. *R Soc Open Sci.* (2014) 1:140216. doi: 10.1098/rsos.140216
- Biau DJ, Jolles BM, Porcher R. P value and the theory of hypothesis testing: an explanation for new researchers. *Clin Orthop Relat Res.* (2010) 468:885–92. doi: 10.1007/s11999-009-1164-4
- Freedman LP, Cockburn IM, Simcoe TS. The economics of reproducibility in preclinical research. *PLoS Biol.* (2015) 13:e1002165. doi: 10.1371/journal.pbio.1002165
- Jager LR, Leek JT. An estimate of the science-wise false discovery rate and application to the top medical literature. *Biostatistics* (2014) 15:1–12. doi: 10.1093/biostatistics/kxt007
- Lindsay DS. Replication in psychological science. *Psychol Sci.* (2015). 26:1827–32. doi: 10.1177/0956797615616374
- Gelman A. Objections to Bayesian statistics. *Bayesian Anal.* (2008) 3:445–9. doi: 10.1214/08-BA318
- McGrayne SB. *The Theory that Would not Die: How Bayes' Rule Cracked the Enigma Code, Hunted Down Russian Submarines, & Emerged Triumphant from Two Centuries of Controversy.* New Haven, CT: Yale University Press (2011).
- Franco A, Malhotra N, Simonovits G. Publication bias in the social sciences: unlocking the file drawer. *Science* (2014) 345:1502–5. doi: 10.1126/science.1255484
- Head ML, Holman L, Lanfear R, Kahn AT, Jennions MD. The extent and consequences of P-Hacking in science. *PLoS Biol.* (2015) 13:e1002106. doi: 10.1371/journal.pbio.1002106
- Aschwanden C. *Science Isn't Broken, Five Thirty Eight.* Available online at: <http://fivethirtyeight.com/features/science-isnt-broken/> (Accessed January 22, 2016), (2015).
- de Winter JC, Dodou D. A surge of p -values between 0.041 and 0.049 in recent decades (but negative results are increasing rapidly too). *PeerJ* (2015) 3:e733. doi: 10.7287/peerj.preprints.447v4
- Simonsohn U. *Posterior-Hacking: Selective Reporting Invalidates Bayesian Results Also.* Philadelphia, PA: University of Pennsylvania, The Wharton School; Draft Paper (2014).
- Coe R. It's the effect size, stupid: what effect size is and why it is important. In: *Paper Presented at the British Educational Research Association Annual Conference.* Exeter (2002).
- Sullivan GM, Feinn R. Using effect size-or why the P value is not enough. *J Grad Med Educ.* (2012) 4:279–82. doi: 10.4300/JGME-D-12-00156.1
- Weissgerber TL, Milic NM, Winham SJ, Garovic VD. Beyond bar and line graphs: time for a new data presentation paradigm. *PLoS Biol.* (2015) 13:e1002128. doi: 10.1371/journal.pbio.1002128
- Peplow M. Social sciences suffer from severe publication bias. *Nature* (2014). doi: 10.1038/nature.2014.15787. Available online at: <http://www.nature.com/news/social-sciences-suffer-from-severe-publication-bias-1.15787>

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Vidgen and Yasseri. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Mechanistic models in computational social science

Petter Holme^{1*} and Fredrik Liljeros²

¹ Department of Energy Science, Sungkyunkwan University, Suwon, South Korea, ² Department of Sociology, Stockholm University, Stockholm, Sweden

Quantitative social science is not only about regression analysis or, in general, data inference. Computer simulations of social mechanisms have an over 60 years long history. They have been used for many different purposes—to test scenarios, to test the consistency of descriptive theories (proof-of-concept models), to explore emergent phenomena, for forecasting, etc... In this essay, we sketch these historical developments, the role of mechanistic models in the social sciences and the influences from the natural and formal sciences. We argue that mechanistic computational models form a natural common ground for social and natural sciences, and look forward to possible future information flow across the social-natural divide.

Keywords: computational social science, mechanistic models, simulation, complex systems, interdisciplinary science

OPEN ACCESS

Edited by:

Yamir Moreno,
University of Zaragoza, Spain

Reviewed by:

Ladislav Kristoufek,
Institute of Information Theory and
Automation (AS CR), Czech Republic
Carlos Gracia-Lázaro,
University of Zaragoza, Spain

*Correspondence:

Petter Holme,
Department of Energy Science,
Sungkyunkwan University,
Suwon 440-746, South Korea
holme@skku.edu

Specialty section:

This article was submitted to
Interdisciplinary Physics,
a section of the journal
Frontiers in Physics

Received: 30 June 2015

Accepted: 01 September 2015

Published: 17 September 2015

Citation:

Holme P and Liljeros F (2015)
Mechanistic models in computational
social science. *Front. Phys.* 3:78.
doi: 10.3389/fphy.2015.00078

Background

In mainstream empirical social science, a result of a study often consists of two conclusions. First, that there is a statistically significant correlation between a variable describing a social phenomenon and a variable thought to explain it. Second, that the correlations with other, more basic, or trivial, variables (called *control*, or *confounding*, variables) are weaker. There has been a trend in recent years to criticize this approach for putting too little emphasis on the mechanisms behind the correlations [1–3]. It is often argued that regression analysis (and the linear, additive models they assume) cannot serve as causal explanations of an open system such as usually studied in social science. A main reason is that, in an empirical study, there is no way of isolating all conceivable mechanisms [4]. Sometimes authors point to natural science as a role model in the quest for mechanistic models. This is somewhat ironic, since many natural sciences, most notably physics, traditionally put more emphasis on the unification of theories and the reduction of hypotheses [1]. In other words, striving to show that two theories could be more simply described as different aspects of a single, unified theory. Rather than being imported from natural or formal sciences, mechanistic modeling has evolved in parallel in the social sciences. Maybe the most clean-cut forms of mechanistic models are those used in computer simulations. Their past, present and future, and the flow of information regarding them across disciplines, are the themes of this paper. Before proceeding, other authors would probably spend considerable amounts of ink to define and discuss central concepts—in our case “mechanism” and “causal.” We think their everyday usage in both natural and social sciences is sufficiently precise for our purpose and recommend [3] to readers with a special interest of details.

In practice, establishing the mechanisms behind a social phenomenon takes much more than simulating a model. Mechanistic models can serve several different purposes en route to establishing a mechanistic explanation. We will make a distinction of *proof-of-concept modeling*, *discovery of hypotheses* and *scenario testing* (described in detail below). There are of course other

ways, perhaps also better, to characterize mechanistic models. These categories are not strict either—they could be overlapping with regard to a specific model. Nevertheless, we think they serve a point in our discussion and that they are fairly well defined.

The idea of proof-of-concept modeling is to test the consistency of a verbal description, or cartoon diagram, of a phenomenon [5]. It is in general hard to make an accurate verbal explanation, especially if it involves connecting different levels of abstraction, such as going from a microscopic to a macroscopic description. A common mistake is to neglect implicit assumptions, some that may even be the convention of a field. With the support of such proof-of-concept models, a verbal argument becomes much stronger. Then one has at least firmly established that the constituents of the theory are sufficient to explain the phenomenon. The individual-based simulations of the Anasazi people (inhabiting parts of the American West millennia ago) by Joshua Epstein, Robert Axtell and colleagues [6] are blueprints of proof-of-concept modeling. In these simulations, the authors combined a multitude of conditions along with anthropological theories to show that they could generate outcomes similar to the archeological records.

The most common use of mechanistic models is our second category—to explore the possible outcomes of a certain situation, and to generate hypotheses. We will see many examples of that in our essay. As a first example, consider Robert Axelrod's computer tournaments to find optimal strategies for the iterated prisoner's dilemma [7]. The prisoner's dilemma captures a situation where an individual can choose whether or not to cooperate with another. If one knows that the encounter is the last one, the rational choice is always not to cooperate. However, if the situation could be repeated an unknown number of times, then it might be better to cooperate. To figure out the way to cope with this situation, Axelrod invited researchers to submit strategies to a round-robin tournament. The winning strategy ("tit-for-tat") was to start cooperating and then do whatever your opponent did the previous step. From this result, Axelrod could make the hypothesis that a tit-for-tat-like behavior is common among both people and animals, either because they often face a prisoner's dilemma or at that such situations, once you face them, tend to be important.

Mechanistic models forecasting social systems are less frequent than our previous two classes. One reason is probably that forecasting open systems is difficult (sometimes probably even impossible) [4]; another that non-mechanistic methods (machine learning, statistical models, etc...) are better for this purpose. A model without any predictive power whatsoever is, of course, not a model at all, and under some conditions all mechanistic models can be used in forecasting, or (perhaps more accurately) scenario testing. One celebrated example is the "World3" simulation popularized by the Club of Rome 1972 book *The Limits to Growth* [8] where an exponentially growing artificial population faced a world of limited resources. Maybe a sign of the time, since several papers from the early 1970s called for "whole Earth simulations" [9, 10]. Echoes of this movement were heard recently with the proposal of a "Living Earth Simulator" [11].

In this essay, we will explore mechanistic models as scientific explanations in the social sciences. We will give an overview of the development of computer simulations of mechanistic models (primarily in the social sciences, but also mentioning relevant developments in the natural sciences), and finally discuss if and how mechanistic models can be a common ground for cross-disciplinary research between the natural and social sciences. We do not address data-driven science in the interface of the natural or social, nor do we try to give a comprehensive survey of mechanistic models in the social sciences. We address anyone interested in using simulation methods familiar to theoretical natural scientists to advance the social sciences.

Influence from the Natural and Formal Sciences

As we will see below, the development and use of computer simulations to understand social mechanisms has happened on quite equal terms as in the natural and formal sciences. It will, however, be helpful for the subsequent discussion to sketch the important developments of computer simulations as mechanistic models in the natural sciences. This is of course a topic that would need several book volumes for a comprehensive coverage—we will just mention what we regard the most important breakthroughs.

The Military Origins

Just like in social science, simulation in natural science has many of its roots in the military from the time around the Second World War. The second major project running on the first programmable computer, ENIAC, started April 1947. The topic, the flow of neutrons in an incipient explosion of a thermonuclear weapon [12], is perhaps of little interest today, but the basic method has never ran out of fashion—it was the first computer program using (pseudo) random numbers, and hence an ancestor of most modern computer simulations. Exactly who invented this method, codenamed Monte Carlo, is somewhat obscure, but it is clear it came out of the development of the hydrogen bomb right after the war. The participants came from the (then recently finished) Manhattan project. Nicholas Metropolis, Stanislaw Ulam and John von Neumann are perhaps most well-known, but also Klara von Neumann, John's wife [12]. It was not only the first program to use random numbers, it was also the first modern program in the sense that it had function calls, and had to be fed into the computer along with the input. As a curiosity, the random number generator in this program worked by squaring eight-digit numbers and using the mid eight digits as output and seed to the next iteration. Far from having the complexity of modern pseudo random number generator (read Mersenne Twister [13]), it gives random numbers of (at least in the authors' opinion) surprisingly good statistical quality.

The first Monte Carlo simulation was not an outright success as a contribution to the nuclear weapons program. Nevertheless, the idea of using random numbers in simulations has not fallen out of fashion ever since, and the Monte Carlo method (nowadays referring to any computational model based on random numbers) has become a mainstay of numerical methods.

Another very significant step for the natural sciences, especially chemistry and statistical physics, by the Los Alamos group was the Metropolis–Hastings algorithm—a method to sample configurations of particles, atoms or molecules according to the Boltzmann distribution (connecting the probability of a configuration and its energy). The radical invention was to choose configurations with a probability proportional to the Boltzmann distribution and weighting them equally, rather than choosing configurations randomly and weighing them by the probability given by the Boltzmann distribution [14]. Hastings name was added to credit his extension of the algorithm to general distributions [15]. Today, this algorithm is an indispensable simulation technique to generate the probability distributions of the state of a system both in natural and social sciences (usually called Markov Chain Monte Carlo, MCMC).

The Monte Carlo project and the MCMC method did not immediately lead to fundamental advances in science itself. Deterministic computational methods, on the other hand, did, and (not surprisingly) post-Manhattan-project researchers were involved. Enrico Fermi, John Pasta, and Stanislaw Ulam (and, like the Monte Carlo project, with undercredited help by a female researcher, Mary Tsingou [16]) studied vibrations of a one-dimensional string with non-linear corrections to Hooke's law (that states that the force needed to extend a spring a certain distance is proportional to the distance). They expected to see the non-linearity transferring energy from one vibrational mode (like the periodic solution of the linear problem) to all other modes (i.e., thermal fluctuations) according to the equipartition theorem [17]. Instead of such a “thermalization” process, they observed the transition to a complex, quasi-periodic state [18] that never lost its memory of the initial condition. The FPU paradox was the starting point of a scientific theme called *non-linear science* that also, as we will see, has left a lasting imprint on social science.

Complexity Theory

Non-linear science has a strong overlap with chaos theory, another set of ideas from natural sciences that influenced social science. Chaos is summarized in the vernacular by the “butterfly effect”—a small change (the flapping of a butterfly's wings) could lead to a big difference (a storm) later. One important early contribution came from Edward Lorenz's computational solutions of equations describing atmospheric convection. He observed that a small change in the initial condition could send the equations off into completely different trajectories [19]. Just like for the FPU paradox, the role of the computational method in chaos theory has largely been to discover hypotheses that later have been corroborated by analytical studies. This line of research has not been directly aimed at discovering new mechanisms; still, ideas and concepts from chaos theory have also reached social sciences [20].

Another natural science development largely fueled by computer simulations, which has influenced social sciences, is that of fractals. Fractals are mathematical objects that embody self-similarity—a river can branch into tributaries, that branch into smaller tributaries, and so on, until the biggest rivers are reduced to the tiniest creeks [21]. At all scales, the branching looks the same. Fractals provide an

analysis tool—the fractal dimension—that can characterize self-similar objects. There are many socioeconomic systems that are self-similar—financial time series [22], the movement of people [23], the fluctuations in the size of organizations [24], etc... Quite frequently, however, authors have not accompanied their measurement of a fractal dimension with a mechanistic explanation of it, which is perhaps why fractals have fallen out of fashion lately.

Fractals are closely related to power-law probability distributions, i.e., the probability of an observable x being proportional to $x^{-\alpha}$, $\alpha > 0$. Power-laws are the only self-similar (or “scale-free”) real-to-real functions in the sense that, if e.g., the wealth distribution of a population is a power law, then a statement like “there are twice as many people with a wealth of 10X than 15X” is true, no matter if X is dollars, euros, yen or kronor [25]. The theories for such power-law phenomena date back to Pareto's lectures on economics published 1896 [26]. Fractals and power-laws are also connected to phase transitions in physics—an idea popularized in Hermann Haken's book *Synergetics* [27].

Next step in our discussion is the studies of artificial life. The central question in this line of research is to mechanistically recreate the fundamental properties of a living system, including self-replication, adaptability, robustness and evolution [28]. The origins of artificial life can be traced to John von Neumann's self-replicating cellular automata. These are configurations of discrete variables confined to an underlying square grid that, following a distinct set of rules, can reproduce, live and die [29]. The field of artificial life later developed in different directions, both toward the more abstract study of cellular automata and to more biology-related questions [28]. It is also strongly linked to the study of adaptive systems (systems able to respond to changes in the environment) [30] and has a few recurring ideas that also are related to social phenomena. The first idea is that simple rules can create complex behavior. The best-known model illustrating this is perhaps Conway's game of life. This is a cellular automaton with the same objectives as that of von Neuman, but with fewer and simpler rules [28]. The second idea (maybe not discovered by the field of artificial life, but at least popularized) is that of *emergence*. This refers to the properties of a system, as a whole, coming from the interaction of a large number of individual subunits. A textbook example is that of murmurations of birds (flocks of hundreds of thousands of e.g., starlings). These can exhibit an undulating motion, fluctuating in density, that in no way could be anticipated from the movement of an individual. Another feature of emergence, exemplified by bird flocks, is that of *decentralization*—there is no leader bird. These topics are common to many disciplines of social science (emergence is similar to the micro-to-macro-transition in sociology and economics). These theories have spawned its own modeling paradigm—agent based models [31–34]—that is similar to what was simply called “simulation” in early computational social science. One first sets up rules for how units (agents) interact with each other and their surroundings. Then one simulates many of them together (typically on a two-dimensional grid) and let them interact. We note that the concept of emergence has also been influential to cognitive, and subsequently behavioral, science. The

idea of cognitive processes being emergent properties of neural networks—connectionism [35]—is nowadays fundamental to our understanding of computational processes in nature [36].

In the 1980's, artificial life, adaptive systems, fractals and chaos were grouped together under the umbrella term complexity science [37]. This was in many ways a social movement gathering researchers of quite marginalized research topics (the Santa Fe Institute, and some similar centers, acted as hubs for this development). Many of the themes within complexity science could probably just as well be categorized as mutually independent fields. This is perhaps best illustrated in that there is no commonly accepted definition of “complexity.” Instead, there are a number of common, occasionally (but not always) connected, themes (like the above-mentioned, emergence, decentralized organization, fractals, chaos, etc...) that together defines the field. On the other hand, there is a common goal among complexity scientists to find general, organizational principles that are not limited to one scientific field. In spirit, this dates back to, at least, von Bertalanffy's general systems theory [38]. The diversity of ideas and applications has not necessarily been a problem for complexity science; on the contrary, it has encouraged many scientists of different backgrounds (including the authors of this paper) to try collaborating, despite the transdisciplinary language barriers.

Game Theory

Game theory is a mathematical modeling framework for situations where the state of an individual is jointly determined by the individual's own decisions and the decisions of others (who all, typically, strive to maximize their own benefit) [39]. Vaccination against infectious diseases is a typical example. If everyone else were vaccinated, the rational choice would be to not get vaccinated. The disease could anyway not spread in the population, whether or not you are vaccinated. Moreover, vaccines can, after all, have side effects, and injections are uncomfortable. If nobody were vaccinated, and the chance of getting the disease times the gravity of the consequences outweighs the above-mentioned inconveniences, then it would be rational to get vaccinated. This situation could, mathematically, be phrased as a minority game [40]. The emergent solution for a population of rational, well-informed and selfish individuals is that a fraction of the agents would get vaccinated and another fraction not. This example is, at the time of writing, the background to a controversy where people getting vaccinated see people resisting vaccination as irresponsible to the society [41].

Game theory has been an especially strong undercurrent in economy and population biology. We note that a special feature of game theory, compared to similarly interdisciplinary theories, is that the various fields using it seem rather well informed about the other fields' progress and not so many concepts have been reinvented. Game theory itself is not a framework for mechanistic models, and especially in population biology (where an individual usually represents a species or a sub-population) it is not clear that is its main use. Nevertheless, there are many mechanistic models in economy and population biology that uses game theory as a fundamental ingredient [42].

Network Theory

Just like complexity and game theory, network theory is a great place for information exchange between the natural and social sciences. Its basic idea is to use networks of vertices, connected pairwise by edges, as a systematic way of simplifying a system. By studying the network structure (roughly speaking, how a network differs from a random network) one can say something about how the system functions as a whole, or the roles of the individual vertices and edges in the system [43, 44]. The multidisciplinary of network theory is reflected in its overlapping terminology—vertices and edges are called nodes and links in computer science, sites and bonds in physics and chemistry, actors and ties in sociology, etc...

Many ideas in network theory originated in social science, and for that reason it may not fit in a section about influences from natural science. Nevertheless, as mentioned, it is a field where ideas frequently flow from the natural and formal sciences to social sciences. Centrality measures like PageRank and HITS were, for example, developed in computer science [43], as were fundamental concepts of temporal network theory (where information about the time when vertices and edges are active is included in the network) [45].

Early Computer Simulations to Understand Social Mechanisms

In this section, we will go through some developments in the use of mechanistic models in social science. We will focus on early studies, assuming the readers largely know the current trends. This is by no means a review (which would need volumes of books), but a few snapshots highlighting some differences and similarities to today's science in the methodologies and the questions asked.

Operations Research

Just like the computer hardware, the research topics for simulation and mechanistic models have many roots in military efforts around the Second World War. Perhaps the main discipline for this type of research is operations research, which is usually classified as a branch of applied mathematics. The objective of operations research is to optimize the management of large-scale organizations—managing supply chains, scheduling crews of ships, planes and trains, etc... The military was not the only such organization that interested the early computer simulation researchers. Harling [46] provides an overview of the state of computer simulations in operation research in the late 1950's. As a typical example, Jennings and Dickins modeled the flow of people and buses in the Port Authority Bus Terminal in New York City during the morning rush hour [47]. They modeled the buses individually and passengers as numbers of exiting, not transferring, individuals. The authors tried to simultaneously optimize the interests of three actors—the bus operators, the passengers, and the Port Authority (operating the terminal). These objectives were mostly not conflicting—in principle it was better for all if the passenger throughput was as high as possible. A further simplifying factor was that the station was the terminus for all buses. The challenge was that

buses stopping to let off passengers could block other buses, thus creating a traffic jam. To solve this problem, the paper evaluated different methods to assign a bus stop to an incoming bus.

Political Science

Although rarely cited today, simulation studies of political decision processes were quite common in the 1950s and 1960s. Crecine [48] reviews some of these models. One difference from today is that these models were less abstract, often focusing on a particular political or juridical organization. The earliest paper we are aware of is Guetzkow's 1959 investigation of the use of computer simulations as a support system for international politics [49]. However, many studies in this field credit de Sola Pool et al.'s simulation of the American presidential elections 1960 and 1964 as the starting point [50]. In their work, the authors gathered a collection of 480 voter profiles that they could use to test different scenarios (with respect to what topics that would turn out to be important for the campaign). Eventually they predicted the outcome of the elections with 82% accuracy.

In their Ph.D. theses, Cherryholmes [51] and Shapiro [52] modeled voting in the House of Representatives by: First, dividing members into classes with respect to how susceptible they were to influence. Second, modeling the influence process via an interaction network where people were more likely to communicate (and thus influence each other) if they were from the same party, state, committee, etc...Cherryholmes and Shapiro also validated their theories against actual voting behavior (something rarely seen in today's simulation studies of opinion spreading [53]). Other authors addressed more theoretical issues of voting systems, such as Arrow's paradox [54, 55] (which states, briefly speaking, that a perfect voting system is impossible for three or more alternatives).

There was also a considerable early interest in simulating decision making within an organization. Apparently the Cuban missile crisis of 1962 was an important source of inspiration. De Sola Pool was, once again, a pioneer in this direction with a simulation of decision-making in a developing, general crisis with incomplete information [56]. Even more explicitly, Smith [57] based his simulation on the personal accounts of the people involved in solving the Cuban missile crisis. Clema and Kirkham proposed yet a model of risks, costs and benefits in political conflicts [58]. Curiously, as late as 2007 there was a paper published on simulating the Cuban missile crisis [59]. However, this paper explores mechanistic modeling as a method of teaching history, rather than the mechanisms of the decision making process itself.

Another type of political science research concerns the evolution of norms. A classic example is Axelrod's 1986 paper [60] where he investigated norms emerging as successful strategies in situations described by game theory. Axelrod let the norms evolve by genetic algorithms (an algorithmic framework for optimization inspired by genetics). In addition to norms, Axelrod also studied metanorms—norms that promote other norms (by e.g., encouraging punishing of people breaking or questioning the norms). Axelrod interpreted the results of the simulation in terms of established social mechanisms supporting

the existence of norms (domination, internalization, deterrence, etc...).

Linguistics

In linguistics, the first computer simulation studies appeared in the mid-1960s. A typical early example is Klein [61] who developed an individual-based simulation platform for the evolution of language. Just like Cherryholmes and Shapiro (above), Klein assumed that the communication was not uniformly random between all pairs of individuals—spouses were more likely to speak to, and learn from, one another, as were parents and children. In multilingual societies, speakers were more likely to communicate to another speaker of the same language (Klein allowed multilingual individuals). A language was represented by a set of explicit grammatical rules (with explicit word classes: nouns, verbs, etc...). Communication reinforced the grammatical rules between the speakers. Klein incremented the time by years and simulated several generations of speakers, but was not entirely happy with the results as communities tended to lose the diversity of their grammar quickly or diverge to mutually incomprehensible grammars. In retrospect, we feel like it was a still a great step forward, where the negative results helped raising important questions about what mechanisms that were missing. More modern models of language evolution have considered much simpler problems [62]. One cannot help thinking that this is to avoid the complexities of reality, and more models in the vein of Klein's 1966 paper would be more important. Later, Klein focused his research on more specific questions like the evolution of Tikopia and Maori [63]. The goal of these early simulation studies was to create something similar to a sociolinguistic fieldwork study. Thus, these were proof-of-concept studies on a more concrete level than today's more theoretically motivated research.

Geography

Demography and geography were also early fields to adopt computer simulations. One notable pioneer was the authors' compatriot Torsten Hägerstrand whose Ph.D. thesis used computer simulations to investigate the diffusion of innovations [64]. His model was similar to two-dimensional disease-spreading models (but probably developed independently of computational epidemiology, where the first paper was published the year before [65]). Hägerstrand used an underlying square grid. People were spread out over the grid according to an empirically measured population distribution. At each iteration of the simulation, there was a contact between two random individuals (where the chance of contact decayed with their separation). If the one of the individuals had adopted the innovation, and the other had not, then the latter would (with 100% probability) adopt it. A goal of Hägerstrand's modeling was to recreate a "nebula shaped" distribution of the innovation (this is further developed in Hägerstrand [66]). To this end, Hägerstrand introduced a concept (still in use) called *mean information field* representing the probability of getting the information (innovation) from the source.

A technically similar topic to information diffusion is that of migration (as in moving one's home). This research dates

back to Ravenstein's 1885 paper "The laws of migration" which is very mechanistically oriented [67]. He listed seven principles for human migration such as: short-distance migration is more common than long-distance; people who migrate far have a tendency to go to a "great centre of commerce or industry." Computer simulation lends itself naturally to exploring the outcomes such mechanisms in terms of demographics. One such example is Porter's migration model where agents were driven by the availability of work and the availability of work was partly driven by where people were. If there was an excess of workers, workers would move to the closest available job opportunity; if there was an excess of vacancies, the closest applicant would be offered the job [68].

The study of human mobility (how people move around both in their everyday lives and extreme situations, such as disasters) is an active field of research. It has even been revitalized lately by the availability of new data sources (see e.g., [23]). One common type of simulation study, involving human mobility data, aims at predicting outbreaks of epidemic diseases. To model potentially contagious contacts between people, one can use more or less realism. However, even for the most realistic and detailed simulations, there is a choice of using the real data to calibrate a model of human mobility [69] or run the simulation on actual mobility data (perhaps with simulations to fill in missing data) [70].

Economics and Management Science

There were many early computational studies in economics that used simulation techniques for scenario testing [71, 72]. A typical question was to investigate the operations of a company at many levels (overlapping with the operations-research section above). Evidently, the researchers saw a future where every aspect of running a business would be modeled—marketing, human resource development, social interaction within the company, the competition with other firms, adoption of new technologies, etc... To make progress, the authors needed to restrict themselves. Birchmore [72], for example, focused on forest firms. Much of his work revolved around a forestry firm's interaction with its resource and the many game theoretical considerations that arise from the conflicting time perspectives of short- and long-time revenues and the competition with other companies. Birchmore only used one or a few combinations of parameter values, rather than investigating the parameter dependence like modern game theory would do. Finally, we note that economics and management science were also early to address questions about validation and other epistemological aspects of computer simulations [73].

Anthropology and Demographics

Anthropology was also early to embrace simulation techniques, especially to problems relating to social structure, kinship and marriage [74]. These are perhaps the traditional problems of anthropology that has the most complex structure of causal explanations, and for that reason are most in need for proof-of-concept-type computer simulations. Gilbert and Hammel [75], for example, addressed the question: "How much, and in what ways, is the rate of patrilineal parallel cousin marriage influenced

by the number of populations involved in the exchange of women, by their size, by their rules of postmarital residence, and by degree of territorially endogamic preference?" To answer these questions, the authors constructed a complex model including villages of explicit sizes, individuals of explicit gender, age and kinship, and rules for how to select a spouse. The model was described primarily in words, in much detail and length. A modern reader would think that pseudocode would make the paper more readable (and certainly much shorter). Probably the anthropology journals of the time were too conservative, or the programming literacy too low, for including pseudocode in the articles.

In a study similar to Gilbert and Hempel, one step closer to demographics, May and Heer [76] used computer simulations to argue that the large family sizes in rural India (of that time) were rational choices for the individuals, rather than a consequence of ignorance and indecision. Around the same time, there were studies of more general questions of human demographics [77], highlighting a transition from mechanistic models for scenario testing to proof-of-concept models and hypothesis discovery.

Cognitive and Behavioral Science

In cognitive science (sometimes bordering to behavioral science), researchers in the 1960s were excited about the prospects of understanding human cognition as a computer program.

Abelson and Carroll [78], for example, proposed that mechanistic simulations could address questions like how a person can reach an understanding ("develop a belief system") of a complex situation in terms of a set of consistent descriptive clauses (encoding, for example causal relationships). Several researchers proposed reverse engineering of human thinking into computer programs as a method to understand cognitive processes [79]. Some even went so far as to interpret dreams as an operating system process [80]. These ideas were not without criticism. Frijda [81] argued that there would always be technical aspects of computer code without a corresponding cognitive function. History seems to give the author right since few studies nowadays pursue replicating human thinking by procedural computer programs. There were of course many other types of studies in this area. For example, early studies in computational neuroscience influenced the behavioral-science side of cognitive science [82].

Sociology

Simulation, in sociology, has always been linked to finding social mechanisms. Even before computer simulations, there were mathematical models for that purpose [83, 84]. As an example of mathematical model building, we briefly mention Nicholas Rashevsky and his program in "mathematical biophysics" at University of Chicago [85, 86]. Trained as a physicist, Rashevsky and his group pioneered the modeling of many social (and biological) phenomena such as social influence [87], how social group structure affect information flow [88], and fundamental properties of social networks [89]. However, Rashevsky and colleagues operated rather disconnected from the rest of academia—mostly publishing in their *Bulletin of Mathematical Biophysics* and often not building on empirical results available.

Perhaps for this reason (even though his contemporaries were aware of his work [90]) is Rashevsky et al.'s direct impact on today's sociology rather limited.

Even though there were stochastic models in sociology in the early 1960's (e.g., [91]), these were analyzed analytically and early sociological computer simulations were off to a rather late start. Coleman [92], Gullahorn and Gullahorn [93, 94] gave the earliest discussions of the prospects of computer modeling in sociology that we are aware of. Coleman discussed both abstract questions about relating social action and social organization, and more concrete ones like using simulation to test social-contagion scenarios of smoking among adolescents. The Gullahorns were more interested in organization and conflict resolution, typically in the interface of sociology and behavioral science. McGinnis [95] presented a stochastic model of social mobility that he analyzed both analytically and by simulations. "Mobility," in McGinnis work, should be read in an extremely general sense, indicating change of an individual's position in any sociometric observable (including physical space).

Markley's 1967 paper on the SIVA model is another early simulation study of a classic sociological problem [96], namely what kind of pairwise relationships that could build up a stable organization. The letters SIVA stands for four aspects of such relationships in an organization facing some situation that could require some action to be taken—Strength (the ratio of how important the two individuals are to the organization), Influence (describing how strongly they influence each other), Volitional (the relative will to act with respect to the situation) and Action (quantifying the joint result of the two actors). These different aspects are coupled and Markley used computer simulations to find fixed points of the dynamics. For many parameter values, it turned out that the SIVA values diverged or fluctuated—which Markley took as an indication that one would not observe such combinations of parameter values in real organizations.

A model touching classical sociological ground that recently has received exceptional amounts of attention is Schelling's segregation model [97]. With this model, Schelling argued that a strong racial segregation (with the United States in mind) does not necessarily mean that people have very strong opinions about the race of their neighbors. Briefly, Schelling spread individuals of two races on a square grid. Some sites were left vacant. Then he picked an individual at random. If this individual had a lower ratio of neighbors of the same race than a threshold value, then he or she moved to a vacant site. It turned out that the segregation (measured as the fraction of links between people of the same race) would always move away from threshold as the iterations converged. Segregation, Schelling concluded, could thus occur without people actively avoiding different races (they just needed to seek similar neighbors), and spatial effects would make a naïve interpretation of the observed mixing overestimating the actual sentiments of the people. The core question—what are the weakest requirements (of tolerance to your neighbors ethnicity) for something (racial segregation) to happen—was a hallmark of Schelling's research and probably an approach that could be fruitful for future studies. We highly recommend Schelling's popular science book *Micromotives and Macrobehavior* [98] as a bridge between the methodologies of natural and social science.

Discussion and Conclusions

The motivation for the use of mechanistic models in social science is often to use them as proof-of concept models. "[I]t forces one to be specific about the variables in interpersonal behavior and the exact relation between them" [93, 99, 100]. The way computer programming forces the researchers to break down the social phenomena into algorithmic blocks helps identifying mechanisms [93, 101]. Other authors point out that with computational methods, the researchers can avoid oversimplifying the problem [50]. Another point of view is that simulation in social sciences is primarily for exploring poorly understood situations and phenomena as a replacement for an actual (in practice impossible to carry out) experiment [48, 102–104]. Such models are obviously closest to hypothesis generators in our above classification. Crane [105] and Ostrom [106] think of computer simulations that, alongside natural languages and mathematics, could describe social sciences. Going a bit off topic, other authors went so far as to using, or recommending to use, computer programs as representations of human cognitive processes [79, 80, 107].

The history of computational studies in social science—as illustrated by our examples—has seen a gradual shift of focus. In the early days, it was, as mentioned, often regarded as a replacement for empirical studies. Such mechanistic models for scenario testing still exists in both natural and social science. However, nowadays it is much more common to use computational methods in theory building—either one uses it to test the completeness of a theoretical framework (proof-of-concept modeling), or to explore the space of possible mechanisms or outcomes (hypothesis discovery).

It is quite remarkable how similar this development has been in the natural and social sciences. At least since mid-1950s, it is hard to say that one side leads the way. This is reflected in how the information flows between disciplines. Looking at the interdisciplinary citation patterns [108] found that out of 203,900 citations from social science journals, 33,891 were to natural science journals, and out of 10,080,078 citations from natural science journals 35,199 were to social science journals. If citations were random, without any within-field bias there would be around 201,000 interdisciplinary citations in both directions, which is 5.9 times the number of social science citations to natural science and 5.7 times the number of natural science citations to social science. In this view, there is almost no inherent asymmetry in the information flow between the areas, only an asymmetry induced by the size difference.

Even though social scientists do not need to collaborate with natural scientists to develop mechanistic modeling, we do encourage collaboration. The usefulness of interdisciplinary collaborations comes from the details of the scientific work. It can help people to see their object system with new eyes. One discipline may, for example, care about the extreme and need input from another to see interesting aspects of the average (cf. phase transitions in the complexity of algorithms [109]). Interdisciplinary information flow could help a discipline overcome technical difficulties. The use of MCMC techniques in the social sciences may be a good example of this. It is,

however, important that such developments come from a need to understand the world around us and not just because they have not been done before.

A major trend at the time of writing is “big data” and “data science.” This essay has intentionally focused on the other side of computational social science—mechanistic models. In practice, these two sides can (and do) influence each other. If it cannot predict real systems at all, a mechanistic model is quite worthless in providing a causal explanation [110, 111]. Modern, large-scale data sets provide plenty opportunities to validate models [112–114]. Another use of big data is in hybrid approaches where one combines a simulation and an empirical dataset, for example simulations of disease spreading on temporal networks of human contacts [45].

As a concluding remark, we want to express our support for social scientists interested in exploring the methods of natural science and natural scientists seeking applications for their methods in the social sciences. To be successful and make most out of such a step, we recommend the social scientist to spend a month to learn a general programming language (Python, Matlab, C, etc...). There is not shortcut (like an integrated modeling environment) to learning the computational subtleties and trade-offs of building a simulation model, and simulation papers often do not mention them. Furthermore, if a social scientist leaves this aspect to a natural scientist,

then she also leaves parts of the social modeling to the natural scientist—collaboration simply works better if the computational fundamentals need not be discussed. To the theoretical natural scientists that are used to simulations, we recommend spending a month reading popular social science books (e.g., [98, 102, 115]). There are too many examples of natural scientists going into social science with the ambition to use the same methods as they are used to—only replacing the natural components by social—and ending up with results that are unverifiable, too general to be interesting, infeasible or already known. While reading, we encourage meditating the following question—why do social scientists ask different questions about society than natural scientists do about nature?

Funding

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2013R1A1A2011947).

Acknowledgments

The authors thank Martin Rosvall for the citation statistics.

References

- Woodward J. Scientific explanation. In: Zalta EN, editor. *The Stanford Encyclopedia of Philosophy*. Available online at: <http://plato.stanford.edu/archives/win2014/entries/scientific-explanation> (2014). doi: 10.1093/acprof:oso/9780195145649.003.0002
- Salmon W. *Causality and Explanation*. Oxford, UK: Oxford University Press (1998). doi: 10.1093/0195108647.001.0001
- Hedström P, Ylikoski P. Causal mechanisms in the social sciences. *Annu Rev Sociol.* (2010) 36:49–67. doi: 10.1146/annurev.soc.012809.102632
- Sayer A. *Realism and Social Science*. New York, NY: Sage (2000). doi: 10.4135/9781446218730
- Servedio MR, Brandvain Y, Dhole S, Fitzpatrick CL, Goldberg EE, Stern CA, et al. Not just a theory: the utility of mathematical models in evolutionary biology. *PLoS Biol.* (2014) 12:e1002017. doi: 10.1371/journal.pbio.1002017
- Epstein JM, Axtell R. *Growing Artificial Societies: Social Science from the Bottom Up*. Washington, DC: Brookings Institution Press (1996).
- Axelrod R. *The Evolution of Cooperation*. New York, NY: Basic Books (1984).
- Meadows DH, Meadows DL, Randers J, Behrens WW. *The Limits to Growth: A Report for the Club of Rome's Project on the Predicament of Mankind*. New York, NY: Universe Books (1972).
- Patterson P. World simulation: a logical extension. *Simulation* (1970) 15:63–4. doi: 10.1177/003754977001500211
- Rau D. World simulation: the need, the feasibility, and a way to start. *Simulation* (1970) 15:64–5. doi: 10.1177/003754977001500212
- Paelucci M, Kossman D, Conte R, Lukowicz P, Argyrakakis P, Blandford A, et al. Towards a living Earth simulator. *Eur Phys J Spec Top.* (2012) 214:77–108. doi: 10.1140/epjst/e2012-01689-8
- Haigh T, Priestley M, Rope C. Los Alamos bets on ENIAC: nuclear Monte Carlo simulations, 1947–1948. *IEEE Ann Hist Comput.* (2014) 36:42–63. doi: 10.1109/MAHC.2014.40
- Matsumoto M, Nishimura T. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans Model Comput Simul.* (1998) 8:3–30. doi: 10.1145/272991.272995
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equations of state calculations by fast computing machines. *J Chem Phys* (1953) 21:1087–92. doi: 10.1063/1.1699114
- Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* (1970) 57:97–109. doi: 10.1093/biomet/57.1.97
- Dauxois T. Fermi, Pasta, Ulam, and a mysterious lady. *Phys Today* (2008) 6:55–7. doi: 10.1063/1.2835154
- Landau LD, Lifshitz EM. *Statistical Physics, Part 1*. 3rd ed. Oxford, UK: Pergamon Press (1980).
- Fermi E, Pasta J, Ulam S. *Studies of Nonlinear Problems. I*. Los Alamos National Laboratory, Los Alamos, NM (1955). Los Alamos Report LA-1940.
- Lorenz EN. Deterministic nonperiodic flow. *J Atmos Sci.* (1963) 20:130–41.
- Kiel LD, Elliott EW (eds.). *Chaos Theory in the Social Sciences: Foundations and Applications*. Ann Arbor, MI: University of Michigan Press (1996).
- Mandelbrot BB. *The Fractal Geometry of Nature*. London, UK: Macmillan (1983).
- Mantegna RN, Stanley HE. *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge, UK: Cambridge University Press (1999). doi: 10.1017/CBO9780511755767
- Brockmann D, Hufnagel L, Geisel T. The scaling laws of human travel. *Nature* (2006) 439:462–5. doi: 10.1038/nature04292
- Mondani H, Holme P, Liljeros F. Fat-tailed fluctuations in the size of organizations: the role of social influence. *PLoS ONE* (2014) 9:e100527. doi: 10.1371/journal.pone.0100527
- Newman MEJ. Power laws, Pareto distributions and Zipf's law. *Contemp Phys.* (2005) 46:323–51. doi: 10.1080/00107510500052444
- Pareto V. *Cours d'Économie Politique. Professe à l'Université de Lausanne*. Lausanne: F. Rouge (1896).
- Haken H. *Synergetik*. Berlin: Springer (1982). doi: 10.1007/978-3-642-96663-7
- Langton CG. *Artificial Life: An Overview*. Cambridge, MA: MIT Press (1998).
- von Neumann J. *Theory of Self-Reproducing Automata*. Urbana, IL: University of Illinois Press (1966).

30. Miller JH, Page SE. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton, NJ: Princeton University Press (2009).
31. Šalamon T. *Design of Agent-Based Models: Developing Computer Simulations for a Better Understanding of Social Processes*. Repin: Bruckner Publishing (2011).
32. Epstein JM. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton, NJ: Princeton University Press (2006).
33. Carley KM, Wallace WA. *Computational Organization Theory*. Berlin: Springer (2001). doi: 10.1007/1-4020-0611-x_143
34. Hedström P, Manzo P. Recent trends in agent-based computational research: a brief introduction. *Sociol Methods Res.* (2015) **44**:179–85. doi: 10.1177/0049124115581211
35. Dawson MRW. *Minds and Machines: Connectionism and Psychological Modeling*. Hoboken, NJ: John Wiley & Sons (2008).
36. Flake GW. *The Computational Beauty of Nature*. Cambridge, MA: MIT Press (1988).
37. Mitchell M. *Complexity: A Guided Tour*. Oxford, UK: Oxford University Press (2011).
38. von Bertalanffy L. *General System Theory: Foundations, Development, Applications*. New York, NY: George Braziller (1968).
39. Hofbauer J, Sigmund K. *Evolutionary Games and Population Dynamics*. Cambridge, UK: Cambridge University Press (1998). doi: 10.1017/CBO9781139173179
40. Challet D, Marsili M, Zhang YC. *Minority Games: Interacting Agents in Financial Markets*. Oxford, UK: Oxford University Press (2005).
41. Honigsbaum M. Balancing unreason: vaccine myths and metaphors. *Lancet* (2015) **385**:763. doi: 10.1016/S0140-6736(15)60423-8
42. Rasmussen E. *Games and Information: An Introduction to Game Theory*. Hoboken, NJ: John Wiley & Sons (1989).
43. Newman MEJ. *Networks: An Introduction*. Oxford, UK: Oxford University Press (2010). doi: 10.1093/acprof:oso/9780199206650.001.0001
44. Barabási A-L. *Network Science*. Cambridge, UK: Cambridge University Press (2015).
45. Holme P, Saramäki J. Temporal networks. *Phys Rep.* (2012) **519**:97–125. doi: 10.1016/j.physrep.2012.03.001
46. Harling J. Simulation techniques in operations research: a review. *Oper Res.* (1958) **6**:307–19. doi: 10.1287/opre.6.3.307
47. Jennings NH, Dickins JH. Computer simulation of peak hour operations in a bus terminal. *Manage Sci.* (1958) **5**:106–20. doi: 10.1287/mnsc.5.1.106
48. Crecine JP. Computer simulation in urban research. *Public Adm Rev.* (1968) **28**:66–77. doi: 10.2307/973586
49. Guetzkow H. A use of simulation in the study of inter-nation relations. *Behav Sci.* (1959) **4**:183–91. doi: 10.1002/bs.3830040302
50. De Sola Pool I, Abelson RP, Popkin S. *Candidates, Issues, and Strategies: A Computer Simulation of the 1960 And 1964 Presidential Elections*. Cambridge, MA: The MIT Press (1965).
51. Cherryholmes CH. *The House of Representatives and Foreign Affairs: A Computer Simulation of Roll Call Voting*. Ph.D. thesis, Northwestern University (1966).
52. Shapiro MJ. *The House and the Federal Role: A Computer Simulation of Roll-Call Voting*. Ph.D. thesis, Northwestern University (1966).
53. Castellano C, Fortunato S, Loreto V. Statistical physics of social dynamics. *Rev Mod Phys.* (2009) **81**:591–646. doi: 10.1103/RevModPhys.81.591
54. Klahr D. A computer simulation of the paradox of voting. *Am Polit Sci Rev.* (1966) **60**:384–90. doi: 10.2307/1953365
55. Tullock G, Campbell CD. Computer simulation of a small voting system. *Econ J.* (1970) **80**:97–104. doi: 10.2307/2230441
56. Kessler AR, de Sola Pool I. Crisiscom: a computer simulation of human information processing during a crisis. *IEEE Trans Syst Sci Cybern.* (1965) **1**:52–58. doi: 10.1109/TSSC.1965.300061
57. Smith RB. Presidential decision-making during the Cuban missile crisis: a computer simulation. *Simul Gaming* (1970) **1**:173–201.
58. Clema J, Kirkham J. CONSIM (Conflict Simulator): risk, cost and benefit in political simulations. In: *ACM '71 Proceedings of the 1971 26th Annual Conference*. New York, NY (1971). p. 226–35. doi: 10.1145/800184.810488
59. Stover WJ. Simulating the Cuban missile crisis: crossing time and space in virtual reality. *Int Stud Perspect.* (2007) **8**:111–20. doi: 10.1111/j.1528-3585.2007.00272.x
60. Axelrod R. An evolutionary approach to norms. *Am Polit Sci Rev.* (1986) **80**:1095–111. doi: 10.2307/1960858
61. Klein S. *Dynamic Simulation of Historical Change in Language Using Monte Carlo Techniques*. Technical Report SP-1908. Santa Monica, CA: System Development Corporation (1964).
62. Perfors A. Simulated evolution of language: a review of the field. *J Artif Soc Soc Simul.* (2002) **5**:4. Available online at: <http://jassoc.soc.surrey.ac.uk/5/2/4.html>
63. Klein S, Kuppin MA, Meives KA. Monte Carlo simulation of language change in Tikopia & Maori. In: *Proceedings of the International Conference on Computational Linguistics*. Stroudsburg, PA (1969). p. 699–729. doi: 10.3115/990403.990424
64. Hägerstrand T. *Innovationsförlöppet ur Korologisk Synpunkt*. Lund: Gleerup (1953).
65. Abbey H. An examination of the Reed-Frost theory of epidemics. *Hum Biol.* (1952) **24**:201–33.
66. Hägerstrand T. A Monte Carlo approach to diffusion. *Arch Eur Sociol.* (1965) **6**:43–67. doi: 10.1017/S0003975600001132
67. Ravenstein EG. The laws of migration. *J Stat Soc Lond* (1885) **48**:167–235. doi: 10.2307/2979181
68. Porter R. Approach to migration through its mechanism. *Geogr Ann.* (1956) **38**:317–43. doi: 10.2307/520255
69. Eubank S, Guclu H, Kumar VSA, Marathe MV, Srinivasan A, Toroczkai Z, et al. Modelling disease outbreaks in realistic urban social networks. *Nature* (2004) **429**:180–4. doi: 10.1038/nature02541
70. Balcan D, Colizza V, Gonçalves B, Hu H, Ramasco JJ, Vespignani A. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc Natl Acad Sci USA.* (2009) **106**:21484–9. doi: 10.1073/pnas.0906910106
71. Cohen KJ. Simulation of the firm. *Am Econ Rev.* (1960) **50**:534–40.
72. Birchmore MJ. *A Review of Planning and Evaluation Models as a Basis for the Simulation of a Forest Firm*. M.Sc. thesis, University of British Columbia (1970).
73. Naylor TH, Finger JM. Verification of computer simulation models. *Manage Sci.* (1967) **14**:92–101. doi: 10.1287/mnsc.14.2.B92
74. Coult AD, Randolph RR. Computer methods for analyzing genealogical space. *Am Anthropol.* (1965) **67**:21–9. doi: 10.1525/aa.1965.67.1.02a00020
75. Gilbert JP, Hammel EA. Computer simulation and analysis of problems in kinship and social structure. *Am Anthropol.* (1966) **68**:71–93. doi: 10.1525/aa.1966.68.1.02a00070
76. May DA, Heer DM. Son survivorship motivation and family size in India: a computer simulation. *Popul Stud.* (1968) **22**:199–210. doi: 10.1080/00324728.1968.10405535
77. Barrett JC. A Monte Carlo simulation of human reproduction. *Genus* (1969) **25**:1–22.
78. Abelson RP, Carroll JD. Computer simulation of individual belief systems. *Am Behav Sci.* (1965) **8**:24–30.
79. Newell A, Simon HA. Computer simulation of human thinking. *Science* (1961) **134**:2011–7. doi: 10.1126/science.134.3495.2011
80. Newman EA, Evans CR. Human dream processes as analogous to computer program clearance. *Nature* (1965) **206**:534. doi: 10.1038/206534a0
81. Frijda NH. The problems of computer simulation. *Behav Sci.* (1967) **12**:59–67. doi: 10.1002/bs.3830120109
82. Green BF. Computer models of cognitive processes. *Psychometrika* (1961) **26**:85–91. doi: 10.1007/BF02289687
83. Edling CR. Mathematics in sociology. *Annu Rev Sociol.* (2002) **28**:197–220. doi: 10.1146/annurev.soc.28.110601.140942
84. Coleman JS. *Introduction to Mathematical Sociology*. New York, NY: Free Press (1964).
85. Cull P. The mathematical biophysics of Nicolas Rashevsky. *BioSystems* (2007) **88**:178–84. doi: 10.1016/j.biosystems.2006.11.003
86. Abraham TH. Nicolas Rashevsky's mathematical biophysics. *J Hist Biol.* (2004) **37**:333–85. doi: 10.1023/B:HIST.0000038267.09413.0d
87. Rashevsky N. Mathematical biology of social behavior. *Bull Math Biophys.* (1949) **11**:105–13. doi: 10.1007/BF02477497

88. Rapoport A. Spread of information through a population with a sociostructural bias: I. Assumption of transitivity. *Bull Math Biophys.* (1953) **15**:523–33. doi: 10.1007/BF02476440
89. Solomonoff R, Rapoport A. Connectivity of random nets. *Bull Math Biophys.* (1951) **13**:107–17. doi: 10.1007/BF02478357
90. Karlsson G. *Social Mechanisms*. Glencoe, IL: The Free Press (1958).
91. White H. Chance models of systems of causal groups. *Sociometry* (1962) **25**:153–72. doi: 10.2307/2785947
92. Coleman JS. The use of electronic computers in the study of social organization. *Eur J Sociol.* (1965) **6**:89–107. doi: 10.1017/S0003975600001156
93. Gullahorn JT, Gullahorn JE. A computer model of elementary social behavior. *Syst Res.* (1963) **8**:354–62. doi: 10.1002/bs.3830080410
94. Gullahorn JT, Gullahorn JE. Some computer applications in social science. *Am Sociol Rev.* (1965) **30**:353–65. doi: 10.2307/2090716
95. McGinnis R. A stochastic model of social mobility. *Am Sociol Rev.* (1968) **33**:712–22. doi: 10.2307/2092882
96. Markley OW. A simulation of the SIVA model of organizational behavior. *Am J Sociol.* (1967) **73**:339–47. doi: 10.1086/224481
97. Schelling TC. Models of segregation. *Am Econ Rev.* (1969) **59**:488–93.
98. Schelling TC. *Micromotives and Macrobehavior*. New York, NY: W. W. Norton (1978).
99. Hare AP. Computer simulations in small groups. *Behav Sci.* (1961) **6**:261–5.
100. Hartman JJ, Walsh JA. Simulation of newspaper readership: an exploration in computer analysis of social data. *Soc Sci Q.* (1969) **49**:840–52.
101. Dutton JM, Briggs WG. Simulation model construction. In: Dutton JM, Starbuck WH, editors. *Computer Simulation of Human Behavior*. New York, NY: John Wiley & Sons (1971). p. 103–26.
102. Simon HA. *The Sciences of the Artificial*. Cambridge, MA: MIT Press (1969).
103. Fleisher A. The uses of simulation. In: Beshers JM, editor. *Computer Methods in the Analysis of Large-Scale Social Systems*. Cambridge, MA: MIT Press (1965). p. 144–6.
104. Naylor TH, Wertz K, Wonnacott TH. Spectral analysis of data generated by simulation experiments with econometric models. *Econometrica* (1969) **37**:333–52. doi: 10.2307/1913541
105. Crane D. Computer simulation: new laboratory for the social sciences. In: Philipson MH, editor. *Automation*. New York, NY: Vintage Press (1962). p. 339–53.
106. Ostrom TM. Computer simulation: the third symbol system. *J Exp Soc Psychol.* (1988) **24**:381–92. doi: 10.1016/0022-1031(88)90027-3
107. Colby KM. Computer simulation of change in personal belief systems. *Syst Res.* (1967) **12**:248–53. doi: 10.1002/bs.3830120310
108. Rosvall M, Bergstrom CT. Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems. *PLoS ONE* (2011) **6**:e18209. doi: 10.1371/journal.pone.0018209
109. Moore C, Mertens S. *The Nature of Computation*. Oxford UK: Oxford University Press (2011). doi: 10.1093/acprof:oso/9780199233212.001.0001
110. Watts DJ. Common sense and sociological explanations. *Am J Sociol.* (2014) **120**:313–50. doi: 10.1086/678271
111. Hindman M. Building better models: prediction, replication, and machine learning in the social sciences. *Ann Am Acad Pol Soc Sci.* (2015) **659**:48–62. doi: 10.1177/0002716215570279
112. Lazer D, Pentland A, Adamic L, Aral S, Barabási A-L, Brewer D, et al. Computational social science. *Science* (2009) **323**:721–3. doi: 10.1126/science.1167742
113. Holme P, Huss M. Understanding and exploiting information spreading and integrating information technologies. *J Comput Sci Technol.* (2011) **26**:829–36. doi: 10.1007/s11390-011-0182-3
114. Pentland A. *Social Physics: How Good Ideas Spread: Lessons From a New Science*. Brunswick: Scribe (2014).
115. Watts DJ. *Everything is Obvious: Once You Know the Answer*. London: Atlantic Books (2012).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Holme and Liljeros. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Citizen Science Practices for Computational Social Science Research: The Conceptualization of Pop-Up Experiments

Oleguer Sagarra¹, Mario Gutiérrez-Roig^{1,2}, Isabelle Bonhoure² and Josep Perelló^{1,2*}

¹ Complexity Lab Barcelona, Departament de Física Fonamental, Universitat de Barcelona, Barcelona, Spain, ² OpenSystems Research, Departament de Física Fonamental, Universitat de Barcelona, Barcelona, Spain

OPEN ACCESS

Edited by:

Javier Borge-Holthoefer,
Qatar Computing Research Institute,
Qatar

Reviewed by:

Raimundo Nogueira Costa Filho,
Universidade Federal do Ceará, Brazil
Gerardo Iñiguez,
Centro de Investigación y Docencia
Económicas A.C., Mexico

*Correspondence:

Josep Perelló
josep.perello@ub.edu

Specialty section:

This article was submitted to
Interdisciplinary Physics,
a section of the journal
Frontiers in Physics

Received: 22 September 2015

Accepted: 09 December 2015

Published: 05 January 2016

Citation:

Sagarra O, Gutiérrez-Roig M,
Bonhoure I and Perelló J (2016)
Citizen Science Practices for
Computational Social Science
Research: The Conceptualization of
Pop-Up Experiments.
Front. Phys. 3:93.
doi: 10.3389/fphy.2015.00093

Under the name of Citizen Science, many innovative practices in which volunteers partner up with scientists to pose and answer real-world questions are growing rapidly worldwide. Citizen Science can furnish ready-made solutions with citizens playing an active role. However, this framework is still far from being well established as a standard tool for computational social science research. Here, we present our experience in bridging gap between computational social science and the philosophy underlying Citizen Science, which in our case has taken the form of what we call “pop-up experiments.” These are non-permanent, highly participatory collective experiments which blend features developed by big data methodologies and behavioral experimental protocols with the ideals of Citizen Science. The main issues to take into account whenever planning experiments of this type are classified, discussed and grouped into three categories: infrastructure, public engagement, and the knowledge return for citizens. We explain the solutions we have implemented, providing practical examples grounded in our own experience in an urban context (Barcelona, Spain). Our aim here is that this work will serve as a guideline for groups willing to adopt and expand such *in vivo* practices and we hope it opens up the debate regarding the possibilities (and also the limitations) that the Citizen Science framework can offer the study of social phenomena.

Keywords: Citizen Science, participation, engagement, computational social science, data, experiments, collective, methods

1. INTRODUCTION

The relationship between knowledge and society has always been an important aspect to consider when one tries to understand how science advances and how research is performed [1, 2]. The general public has, however, mostly been left out of this methodology and creation processes [3, 4]. Citizens are generally considered as passive subjects to whom only finished results are presented in the form of simplified statements; yet paradoxically, we implicitly ask them to support and encourage research. The acknowledgment of this *ivory tower* problem has recently opened up new and exciting opportunities to open-minded scientists. The advent of digital communication technologies, mobile devices and Web 2.0 is fostering a new kind of relation between professional scientists and dedicated volunteers or participants.

Under the name of Citizen Science (CS), many innovative practices in which “volunteers partner with scientists to answer and pose real-world questions” (as stated in the Cornell Ornithology Lab web page; one of the precursors of CS practices in the 1980s) are growing rapidly worldwide [5–8]. Recently, CS has been formally defined by the Societize White Paper as: “general public engagement in scientific research activities when citizens actively contribute to science either with their intellectual effort or surrounding knowledge or with their own tools and resources” [9]. This open, networked and transdisciplinary scenario, favors more democratic research, thanks to contributions from amateur or non-professional scientists [10]. Over the last few years, important results have been published in high-impact journals by using participatory practices [6, 7]. All too often the hidden power of thousands of hands working together is making itself apparent in many fields, and showing its performance to be comparable to (or even better than) expensive supercomputers when used to analyse/classify astronomical images [11], to reconstruct 3D brain maps based on 2D images [12], or to find stable biomolecular structures [13], to name very few of the cases with a large impact. Citizen contributions can also have a direct impact on society by, for instance, helping to create exhaustive and shared geolocalized datasets [14] at a density level unattainable by the vast majority of private sensor networks (and at a much reduced cost) or by collectively gathering empirical evidence to force public administration action (for example, the shutdown of a noisy factory located in London [15]). Most active volunteers can contribute by providing experimental data to widen the reach of researchers, raise new questions and co-create a new scientific culture [3, 16].

Computational social science (CSS) is a multidisciplinary field at the intersection of social, computational and complexity sciences, whose subject of study is human interactions and society itself [17, 18]. However, CS practices remain vastly unexplored in this context when compared to other fields such as environmental sciences, in which they already have a long history [19–21]. Attempts to incorporate the participation of ordinary citizens as playing an important role can be found in fields such as experimental economics [22], the design of financial trading floors [23], and human mobility [24]. Work on the emergence of cooperation [25] and the dynamics of social interactions [26] is also noteworthy. All these experiments yielded important scientific outcomes, with protocols that are well-established and robust within the behavioral sciences (see for instance [27] and [28]), but unfortunately they remain on the very first level of the CS scale [15]. At that level, citizens are involved only as sensors or volunteer subjects for certain experiments in strictly controlled environments; their participation and potential are only partially unleashed. One possible way out of this first level was already provided by Latour [3, 29], when he proposed collective experiments in which the public becomes a driving force of the research. Researchers *in the wild* are then directly concerned with the knowledge they produce, because they are both objects and subjects of their research [4]. Some interesting research initiatives have emerged along these lines and involve massive experiments in collaboration with a CS foundation, such

as Ibercivis [30] or through online platforms such as *Volunteer Science* (a Lazer Lab platform). More radical initiatives consider collaboration with artists as well, and some have been realized in museums or exhibitions and as large-scale performance art [31, 32].

CSS research has also recently been applied in the so-called “big data” paradigm [33, 34]. Much has been said about it and the possibilities it offers to society, industry and researchers. “Smart cities” pack urban areas with all kinds of sensors and integrate the information into a broad collection of datasets. Mobile devices also represent a powerful tool to monitor real-time user-related statistics, such as health, and major businesses opportunities are already being foreseen by companies. However, these approaches again treat citizens as passive subjects from whom one records private data in a non-consensual way, and throws up the aggravated problem that the *unaware* producers of these data (i.e., citizens) lose control of their use, exploitation and analysis. The validity of the conclusions drawn from the analysis of such datasets are still today a subject of discussion, mainly due to poor control of the process of gathering the data (by the public in general and by scientists in particular), inherent population and sampling biases [35] and the lack of reproducibility, among other systemic problems [36]. Last but not least, the big data paradigm has so far failed to provide society with the necessary public debate and transparent practices, adopting the *bottom-up* approach it preconizes. It currently relies on huge infrastructures only available to private corporations, whose objectives may not coincide with those of researchers and the citizenry, and provides conditioned access to the data contents which, in addition, generally cannot be freely (re)used without filtering.

Our purpose here, however, is not to discuss problems inherent to big data. Rather, the approach we present aims to explore the potential of blending interesting features recently developed by big data methodologies with the ambitious and democratic ideals of CS. Public participation and scientific empowerment induce a level of (conscious) proximity with the subjects of the experiments that can be a highly valuable source of high-quality data [37, 38], or at least, of non-conflictive information with regards to data anonymity [39], that may correct biases and systematic experimental errors. This approach is potentially a way to overcome privacy and ethical issues that arise when collecting data from digital social platforms, while keeping high standards of participation [33, 40, 41]. Moreover, CS projects can use a vast variety of social platforms to optimize dissemination, encourage and increase participation and develop *gamification* strategies [42] to reinforce engagement. The so-called Science of Citizen Science studies the emergent participatory dynamics in this class of projects [43, 44], so that this also opens the door to new contexts within which study social phenomena.

The open philosophy at the heart of CS methods, such as open data licensing and coding, can also clearly improve science–society–policy interactions in a democratic and transparent way [45] through so-called deliberative democracy [46]. The CS approach simultaneously represents a powerful example of responsible research and innovation (RRI) practices included in the EU Horizon 2020 research programme [47] and

the Quadruple Helix model in which government, industry, academia and civil participants work together to co-create the future and drive structural changes far beyond the scope of what any one organization or person could do alone [48]. Along these lines, we consider that the potential of CSS when adopting CS methods is vast, since its subject of study is citizens themselves. Therefore, their engagement with projects that study their own behavior is highly likely, since it has an immediate impact on their daily lives. As a result, large motivated communities and scientists can work hand in hand to tackle the challenges arising from CSS, but also collectively circumvent potential side effects. The possibility of reaching wider and more diverse communities will help in the refinement of more universal statements avoiding the population biases [49] and problems of reproducibility present in empirical social science studies [50, 51]. Another important advantage of working jointly with different communities is that it allows scientists to set up *lab-in-the-field* or *in vivo* experiments, which instead of isolating subjects from their natural urban environment—where socialization takes place—are transparent, fully consensual and enriched, thanks to the active participation of citizens [4, 31]. Such practices and methodologies, however, are still far from being well established as a standard tool for CSS research.

The main goal of this paper is precisely to motivate the somewhat unexplored incorporation of CS practices into CSS research activities. Given the arduous nature of such a task, here, we limit ourselves to a *reformatting* of existing standard experimental strategies and methods in science through what we call “pop-up experiments” (PUEs). Such a concept has been shaped by the lessons we have learned while running experiments in public spaces in the city of Barcelona (Spain). Section 2 introduces this very flexible solution which makes collective experimentation possible, and discusses its three essential ingredients: adaptable infrastructure, public engagement and the knowledge return for citizens. Finally, Section 3 concludes the manuscript with a discussion of what we have presented to that point, together with some considerations concerning the future of CS practices within CSS research.

2. A FLEXIBLE SOLUTION FOR CITIZEN SCIENCE PRACTICES WITHIN CSS: THE POP-UP EXPERIMENT

2.1. Context and Motivation

Over the last 4 years, the local authorities in Barcelona (i.e., City Council) and its Creativity and Innovation Direction in collaboration with several organizations have set as an objective the exploration of the possibilities of transforming the city into a public *living lab* [48], where new creative technologies can be tested and new knowledge can be constructed collectively. This has been done through the Barcelona Lab platform and one of its most pre-eminent actions has been establishing CS practices in and with the city. The first task was to create the Barcelona Citizen Science Office and to build a community of practitioners where most of the CS projects from different research institutions in Barcelona could converge. The Office serves as a *meeting point*

for CS projects, where researchers can pool forces, experiences and knowledge, and also where citizens can connect with these initiatives easily and effectively. The second task is directly linked with the subject of this paper and was conceived to test how far the different public administrations can go in opening up their resources to collectively run scientific experiments [31]. The CS toolbox clearly provides the perfect framework for the design of public experiments, and exploration of the emergent tensions and problematic issues when running public living labs. Furthermore, the involvement of the City Council provided us with the opportunity to embed these experiments into important massive cultural events, which constitute the perfect environment for reinforcing the openness and transparency of our research process with respect to society, or at least to the citizens of Barcelona.

We have conducted several experiments to put CS ideals into practice and test their potential in urban contexts. In contrast with existing environmental CS projects in other cities such as London or New York (with civic initiatives such as *Mapping for Change* or *Public Lab*, respectively), we focussed our attention on CSS related problems. Our aim was to explore relations between city, citizens and scientists, which we considered had been neglected or inadequately addressed. More specifically, this rather wild testing consisted of seven different experiments performed between 2012 and 2015 on three different topics that address different questions: human mobility (How do we move?), social dilemmas (How cooperative are we?) and decision-making process (How do we take decisions in a very uncertain environment such as financial markets?). These are summarized in **Table 1** and fully described in Section 4. Other points common to these experiments are the large number of volunteers that participate (up to 541 in a single experiment; 1255 in total) and the consequent large number of records (up to 18,525 decisions for a single experiment, 55,390 entries in total), despite the rather limited budget allocated to the experiments (from 1000 EUR to 4000 EUR per experiment, and around 2200 EUR on average). Despite dealing with different research questions, common problems (related to CS practices, but also to CSS research studying human behavior in general [34]) were identified and potential solutions were developed to overcome them in all cases. Some of the potential solutions that were implemented were successful, others were not; but all the experiences have shaped the concept and the process of experimentation in CSS research consistent with CS ideals.

2.2. Definition of a Pop-Up Experiment and the Underlying Process

The generic definition of a *pop-up*, according to the Cambridge dictionary, is: “Pop-up (adj.): used to describe a shop, restaurant, etc. that operated temporarily and only for a short period when it is likely to get a lot of customers.” From the initial stages, we thought that this description fitted well into our non-permanent but highly participatory experimental set-up when applying CS principles to CSS research in urban contexts. The parallel is very illustrative to understand better a much more formal definition that we use to describe our approach from a

TABLE 1 | Summary of our seven experiments to test CS practices with pop-up experiments in urban contexts and through computational social science research.

Experiments	Bee-Path(I)	Bee-Path(2)	Cooperation(I)	Cooperation(2)	Mr. Banks(I)	Mr. Banks(2)	Dr. Brain
GENERAL INFORMATION							
Date	June 2012	June 2013	December 2012	March 2014	December 2013	July 2015	December 2014
Topic	Human mobility	Human mobility	Social dilemmas	Social dilemmas	Decision making	Decision making	Social dilemmas
Number of volunteers	101 (27 valid)	68	168	52	283	42	541
Number of records	10,312 GPS points	10,315 GPS points	4,200 decisions	1,300 decisions	18,525 decisions	2,372 decisions	8,366 decisions
Budget (approx.)	4000 EUR	3500 EUR	2000 EUR	1000 EUR	2500 EUR	1300 EUR	1000 EUR
EXPERIMENTAL INFRASTRUCTURE AND SET-UP							
Hosting Event	Science fair	Science fair	Games festival	Secondary school	Games festival	Conference	Games festival
Space (m ²)	Outdoor booth (9 m ²)	Outdoor booth (18 m ²)	Room (20 m ²)	4 Classrooms (200 m ²)	Prepared space (80 m ²)	Prepared space (20 m ²)	Prepared space (60 m ²)
Duration	1.5 day (weekend)	1.5 day (weekend)	1.5 day (weekend)	0.25 day	2 days (weekend)	2 days	2 days (weekend)
Typology of volunteers	Science fair visitors (free entrance)	Science fair visitors (free entrance)	Festival visitors (free entrance)	Schoolchildren, 12–13 years old	Festival visitors (free entrance)	Conference participants	Festival visitors (free entrance)
Recording Device	Mobile phones	Mobile phones	Laptops	Laptops	Tablets	Tablets	Tablets
PUBLIC ENGAGEMENT TOOLS							
Number of scientists and facilitators	8	7	7	4	6	4	10
Preparatory actions	Official event programme Web page Press Release Logo Design	Official event programme Web page	Official event programme	–	Official event programme Character and platform creation Press release Targeted mailing	Official event programme Character and platform creation Press release Targeted mailing Twitter campaign	Official event programme Character and platform creation Press release Targeted mailing Twitter campaign
On-site visibility	Location at the entrance Panels Flyers	Location far from the entrance Panels Flyers	Isolated room Panels Actor	–	Location in the main hall Panels Flyers	Location at the main entrance Panels Flyers	Location on main entrance path Panels Flyers
On-site promotion tools	Screen (real-time path visualization)	Screen (real-time path visualization) Actors	–	–	Screen with players ranking Actor Merchandising	Merchandising Actor Merchandising Scenography	Screen with announcement Actor Merchandising Scenography
OUTCOME AND RETURN FOR THE PUBLIC							
Scientific paper	Submitted	None	Published	Published	Submitted	Submitted	Submitted
Sci. Communications	2	None	4	4, same as Coop(1)	2	In preparation	In preparation
Website	bee-path.net	bee-path.net	nectunt.bifi.es	nectunt.bifi.es	mr-banks.net	mr-banks.net	nectunt.bifi.es
OpenData	Data and software	Data and software	–	–	Data available	Data available	In preparation

(Continued)

TABLE 1 | Continued

Experiments	Bee-Path(I)	Bee-Path(2)	Cooperation(I)	Cooperation(2)	Mr. Banks(I)	Mr. Banks(2)	Dr. Brain
Short-term return	On-site discussion with scientists and data visualization	On-site discussion with scientists and data visualization	On-site discussion with scientists Economic reward	Coupons for books and office material	On-site discussion with scientists Ranking visualization	On-site discussion with scientists Ranking visualization	On-site discussion with scientists Coupons for a prize
Medium-term return	Personalized report sent by email	Personalized report sent by email	-	Letter to parents to school	-	-	Personalized report sent by email
Long-term return	Public event to present the results and Courses	Public event to present the results and Courses	Targeted mailing about scientific publication and results	Targeted mailing about scientific publication and results	In preparation	In preparation	In preparation

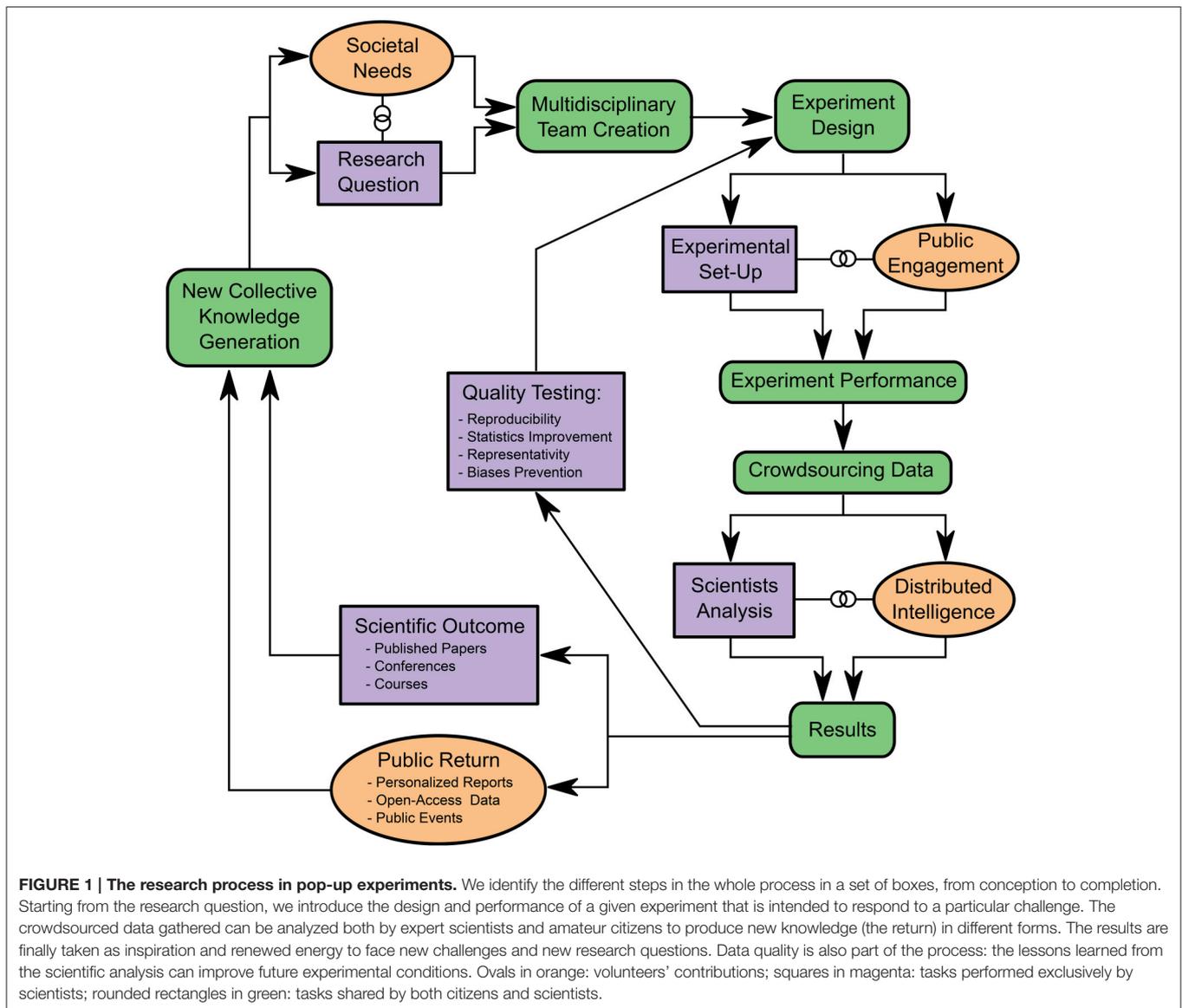
The first column contains the different aspects that each of the experiments considered in terms of experimental infrastructure and set-up; public engagement tools and outcome; and the return for the public. The other columns specify how the different experiments consider these aspects. Further details are provided in the Section 4, while Section 2 discuss different aspects in more depth.

theoretical perspective. It is based on the expertise gained from the seven experiments carried out over the last 4 years and reads:

A PUE is a physical, light, very flexible, highly adaptable, reproducible, transportable, tuneable, collective, participatory and public experimental set-up for urban contexts that: (1) applies Citizen Science practices and ideals to provide ground-breaking knowledge; and (2) transforms the experiment into a valuable, socially respectful, consented and transparent experience for non-expert volunteer participants with the possibility of building common urban knowledge that arises from fact-based effective knowledge valid for both cities and citizens.

In our case, we apply this concept to CSS with the aim of answering very specific research questions with the participation of larger population samples than those in behavioral experiments. The research process that emerges from PUEs can be synthesized in the flow diagram in **Figure 1**. The whole process starts with a research question or a challenge for society that may be promoted by citizens or scientists, but also by private organizations, public institutions or civil movements. The initial impulse helps to create an adequate research group, which will need to be multidisciplinary if it is to tackle a complex problem consisting of many intertwined issues. The group then co-creates the experiment both considering the experimental set-up and the tasks that unavoidably involve public engagement. The experiment is then carried out and data are generated collectively (crowdsourced) under the particular constraints of public spaces which depend not only on the conditions designed by the scientists but also on many other practical limitations. The data are then analyzed using standard scientific methods, but non-professional scientists are also invited to contribute to specific tasks [11, 12] or by using other non-standard strategies in the exploration of the data [13]. These two contributions by volunteers make up what we call “distributed intelligence” and generate results that it is difficult to match by conventional computer analysis. The results can take many forms, depending on the audience being addressed; from a scientific paper to personalized reports that can be read by any citizen or even recommendations that are valid for policymakers at the city level. Finally, the whole process can generate the impulse necessary to promote and face a new social and scientific challenge or an existing need through the same scheme.

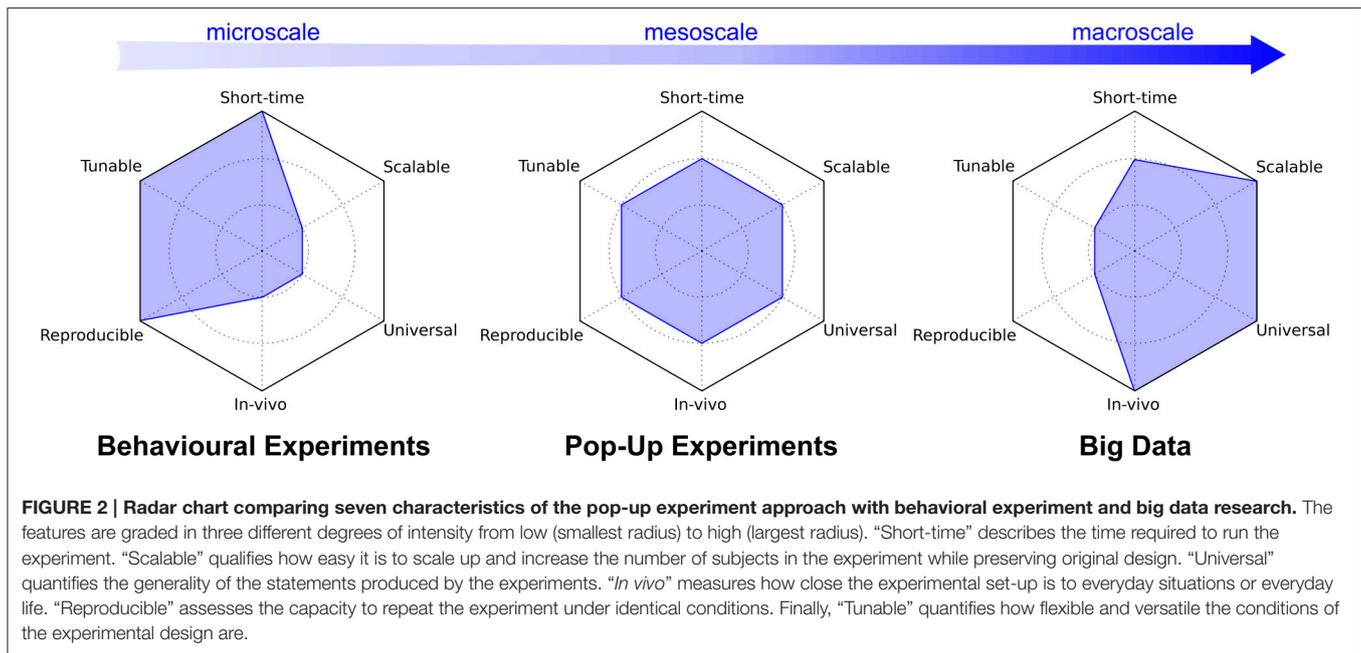
The PUE solution also represents middle ground between behavioral science experiments and big data methodologies. To understand the context in which the PUEs we propose can be placed better, **Figure 2** compares the different approaches considered using a radar chart that qualitatively measures, with three degrees of intensity (low, medium and high), six different aspects that characterize each type of experiment. We can observe that the three different approaches cover different areas. Behavioral experiments and big data have a limited overlap, while the PUEs share several aspects with the former two. One might argue that the excess of openness of CS constitutes a severe limitation with respect to objectivity, compared to the solid experimental protocols in behavioral science [27, 52]. However, it is also true that the highly participatory nature of



CS can be very effective at reaching a more realistic spectrum of the population and a larger sample thereby obtaining more general statements with stronger statistical support (see [28], for alternative and complementary methods). Since it is directly attached to real-world situations, the PUE solution avoids the danger of exclusive and distorted spaces of *in vitro* (or *ex vivo*) laboratory experiments. It also brings additional values to the more classic social science lab-in-the-field experiments which generally limit interaction among subjects and scientists as much as possible. At the other extreme, CS practices will never be able to compete in terms of the quantity of data with the big data world, but this can be compensated for. A better understanding of the volunteers involved and improved knowledge of their peculiarities helps to avoid possible biases. Furthermore, the active nature of PUEs allows some conditions of the experiments to be tuned to explore alternative scenarios.

PUEs can indeed be an alternative to the controversial virtual labs in social networks and mobile games which have yielded interesting results, for instance, in emotional contagion with experiments on the Facebook platform [53], not without an intense public debate on ethical and privacy issues concerning the way the experiments were performed [54].

We think that PUEs could become an essential approach for the empirically testing of the many statements of CSS which is complementary to the lab-in-the-field, virtual labs and *in vitro* experiences. For this to happen, we have identified the main obstacles that hinder the development of CS initiatives with respect to other forms of social experimentation. They can be grouped into three categories: infrastructure, engagement, and return. In what follows, we detail each of the obstacles and illustrate the solutions that PUEs offer, together with practical examples applied to each case.



2.3. Light and Flexible Experimental Infrastructure

By infrastructure, we understand all the logistics necessary to make the experiments possible. In a broad CS context, the necessary elements differ from those of orthodox scientific infrastructure. As discussed in Bonney et al. [6] and Franzoni and Sauermann [55], they include other tools, other technical support and other spaces. The second block of **Table 1** lists some of the elements we have deemed capital to satisfactorily collect reliable data. PUEs should be designed favoring scalability, in the sense of easily allowing an increase in the population sample size or repetition of the experiment in another space. To make this possible, the experiments must rely on solid and well-tested infrastructure, with an appealing volunteer experience to avoid frustrating the participants. When considering the experimental set-up, we used several strategies to foster participation and ensure the success of the experiments.

First, we physically set up the PUEs in very particular contexts in urban areas and, in all cases, we placed them in crowded (moderately to highly dense) places, to reach volunteers easily. In other words, we preferred to go where citizens were instead of encouraging them to come to our labs. To make this possible, the City Council offered specific windows, for instance, at a couple of festivals, as hosting events (the Bee-Path(1), Bee-Path(2), Cooperation(1), Mr. Banks and Dr. Brain experiments). This meant that we had to adapt to these specific out-of-the-lab and *in vivo* contexts; the logistics and the composition of the research teams thus unavoidably became more diverse, complex, heterodox and highly multidisciplinary. In collaboration with the event organizers, we then prepared a specific space, of reduced dimensions, for the experiment, where the volunteers (whose typology was different in each case) could participate through a recording device.

Second, PUEs demand that the devices used by participants to collect and manipulate data, either actively or passively, must be familiar to them. In our experiments we designed specific software to run on laptops (the Cooperation(1) and Cooperation(2) experiments), mobile phones (the Bee-Path(1) and Bee-Path(2) experiments) and tablets (the Mr. Banks and Dr. Brain experiments). However, it may also be possible to use cameras, video cameras, or any other sophisticated device as long as the participants easily become familiarized with it after a few instructions or a tutorial. This sort of infrastructure is in the end what allows us to carry out experiments in a participant’s everyday (not strange: *in vivo*) environment. Initially, we overlooked this aspect in the design of the set-up and the allocation of resources, but having a user-friendly interface is important if our aim is for people to behave *normally*. Similarly, both the instructions and interface should be understandable and manageable for people of all ages.

Third, in order to study social behavior in different environments, the experiments need to be adaptable, tunable, transportable, versatile and easy to set up in different places. All the devices mentioned above fulfill this requirement as well.

Fourth, PUEs are typically one-shot affairs, since they are hosted at a festival, a fair or in a classroom, which means that they are concentrated in time (with a duration of 0.25–2 days) and there is no chance of a second shot. All the collectable data could be threatened if something goes wrong. Extensive beta testing and defensive programming is imperative to ensure that the collected data is reliable. It is also necessary to anticipate potential problems: one must be flexible enough to be able to retrace alternative research questions on the fly if the PUE location and conditions are not fully satisfactory to respond to the initial research purposes.

Finally, numbers matter and experiments must reach enough statistical strength for rigorous analysis to be performed, and this needs to be carefully taken into account during the design phase of the experiment (see **Figure 1**). Typically, the more expensive devices are, the fewer data collectors you can have; so the capacity to collect data is affected and this is not a very effective strategy in a rather short lived one-shot event. Therefore, cheap infrastructure favors scalability in the end. Alternatively, new collaborators need to be found or an extra effort is required to find sponsorship (for instance, related to science outreach) which, in any case, will complicate the preparation phase of the experiment. Scalability is indeed interesting, but it has its side effects as well: relying on infrastructure provided by volunteers, such as smart phones, can greatly influence the quality of the data and its normalization (one of the central problems related to the big data paradigm). In the Bee-Path(1) experiment, where we used the GPS of the smartphones of the participants themselves, the cleaning process was far more laborious than in the other cases, due to this (see numbers in **Table 1**). In contrast, the Cooperation(1), Cooperation(2), Mr. Banks and Dr. Brain experiments, where we designed and programmed the software and supplied the hardware for the experiences of the participants, did not require much post-collection treatment.

2.4. Public Engagement Tools and Strategies

PUEs are physically based and rooted in particular, temporal, local contexts. This delimited framework allows dissemination resources and efforts to be concentrated at a given spot over a certain time, which increases the effectiveness and efficiency of the campaign in terms of both workforce and budget. Additionally, the one-shot nature of PUEs allows us to avoid the problem of keeping participants engaged in an activity spanning long periods of time; but consequently they rely completely on constantly renewing the base of participants (which may require higher dissemination). To this end, the initial action was the creation of a census of volunteers shared by all members of the Barcelona Citizen Science Office research group.

Another factor of major importance is related to the contact between researchers, organizers and citizens in the set-up of the experiment. This allows for pleasant dialogue and exchange of views that in turn helps to frame the scientific question being studied, as well as developing possible improvements for future experiences. This can be achieved by stimulating the curiosity of participants concerning the experiment and the research associated with it. The research question should be focussed and understandable to an average person who is not an expert in the field.

To engage citizens in such a dialogue, however, requires certain steps. First, it is necessary to attract potential participants with an appealing set-up. This includes location in the physical space, but also an effective publicity campaign on the days prior to the experiment (see preparatory actions in **Table 1**). It is important to offer a harmonized design with common themes that citizens can relate to the experiment. To make our material appealing, we collaborated with an artist (the Bee-Path(1) and Bee-Path(2) experiments) and a graphic designer (the

Cooperation(1), Cooperation(2), Mr. Banks(1), Mr. Banks(2), and Dr. Brain experiments) whose main contribution was the creation of characters associated with each experiment. The function of these characters, not far from to the world of cartoons, was to attract the attention of the public, but also to present the experiment in most of the cases as an attractive game, since one of the most powerful elements that engages people in an activity is the expectation of having fun. It is certainly possible to maintain scientific rigor while using *gamification* strategies to create an atmosphere of play for the study, thus transforming it into a more complete experience [56]. Moreover, actors were used as human representations of these characters (Mr. Banks and Dr. Brain). The actors were indeed an important element to bring on-site attractiveness to the experiment, along with: having a large team of scientists and facilitators present (a rotating team of up to 10 people); an optimum and visible location inside the event space; and material/devices to promote the experiment, such as screens to visualize in real time the results of the experiments or promotional material (flyers and merchandise). Based on the experience, we have optimized all these ingredients and we have included some scenographic elements in the last experiment performed within the Barcelona DAU Board Games Festival (Dr. Brain).

Notwithstanding, simply getting a large number of people engaged is not enough. Additionally, it is also necessary to aim for universality in the population sample. The experiment must be designed in such a way that people of all ages and conditions can really participate. Furthermore, a PUE has to be transportable at a minimal cost after it has been implemented once, in order to be reproduced in different environments (which may favor certain types of population). As an example, the Cooperation(1) and Cooperation(2) experiments are very illustrative. In the Cooperation(1) experiment, we discovered that different age groups, especially children ranging from 10 to 16 years old, behaved in different ways and cooperated with different probabilities. Apparently, children were more volatile and less prone to cooperate than the control group in a repeated *prisoner's dilemma*. Fifteen months later, we repeated the experiment in a secondary school [12–13 years old: the Cooperation(2) experiment]. On one hand, in this case the results showed the same levels of cooperation as the control group in Cooperation(1) experiment. However, on the other hand, same volatile behavior, more intense than for the control group in Cooperation(1), was again observed. Therefore, thanks to the repetition of the experiment, we rejected the early idea of different levels of cooperation in children, while at the same time it strengthened the claim that children exhibit volatile behavior (Poncela-Casasnovas et al., submitted).

2.5. Outcome and Return for the Public

Last but not least come the factors related to the management of the aftermath of the experiment (fourth section of **Table 1**). PUEs, as we have implemented them, are intrinsically cross-disciplinary and involve a large number of agents and institutions, which in turn may have diverging interests and expectations regarding the outcome of a particular experiment.

Any successful PUE must be able to accommodate all these interests and create positive environments of collaboration in which all the actors contribute in a mutually cooperative way.

The organizers of a festival will, for instance, find in PUEs an innovative format with participatory activities to add to their programme [the Cooperation(1), Mr. Banks and Dr. Brain experiments]. PUEs can also be a transparent and proactive system of gathering data to provide information and opportunities for analysis of the event itself: useful for planning and improvement. The Bee-Path(1) experiment studied how visitors moved around a given space and provided information based on actual facts that could be used to improve the spatial distribution in future editions of the fair where it was developed. City, local and other administrations will find in PUEs an innovative way to establish direct contact with citizens; to co-create new knowledge valid for the interests of the city as a whole and eventually to generate a census of highly motivated citizens, prepared to participate in this kind of activity. Scientists will obviously try to publish new research based on the data gathered.

All these expectations are very different from each other and should converge organically if a collective experiment is to be run successfully [3]. However, we should not forget to include quite specifically the expected return for our central actors: the volunteers. Their contribution is essential in CS practices [10] and it is therefore completely fair to argue that citizens who agree to participate in these initiatives need to see a clear benefit from their perspective, comparable to (albeit different from) that of the local authorities, festival organizers, scientists or any other contributors. Moreover, the high degree of concentration in space and time, together with the intense public exposure of PUEs, increase volunteer expectations even more compared to other ordinary cases in CS [11, 12]. The face-to-face relationship established between researchers and citizens in all the PUEs we have run testifies to this being a very delicate issue that needs to be managed with great care.

Any PUE should manage expectations on three different time scales: short term (during and immediately after the experiment), medium term (a week or a month afterwards), and long term (the following months or even years). In some of the experiments we failed in this aspect on at least one of the three time scales, since we did not properly anticipate the effort required to respond to the expectations of the volunteers.

The short term responds to basic curiosity. This point is related to engagement and the experimental set-up: the physical presence of scientists (with no mediation) allows them to explain the experiment in the most convinced and convincing way, and thus to motivate people. Also, the introduction of large screens where the progress of other participants can be followed in real time helped in this matter. In the Bee-Path experiments [Bee-Path(1) and Bee-Path(2)] we showed the GPS locations of the participants on a map; while in the Mr. Banks experiment we showed a ranking of the best players (best performances by the participants). This information was intended to boost participation and it was also chosen in such a way that it distorted the questions addressed to a minimally degree and thus did not influence the results of each experiment. The medium term

relates to expectations regarding the results of the experiment. Participants want to know whether the set-up was successful and whether they performed *well enough*. In order to complement their short-term experience, it is important to keep participants informed as to the outcomes of the PUE. An example of a medium timescale is the Dr. Brain case, where a personalized report of performance during the experiment was sent to each participant by e-mail. In some cases, and based on these results, this also generated new dialogue between scientists and citizens. The last timescale to be managed consists of a more formal way of presenting the results of the study, through public presentations and talks. So, in the case of the Bee-Path(1) experiment, outreach conferences, public debates and even a summer course for (graduate and undergraduate) students interested in CS practices were organized.

All these are important for the success of PUEs and should be clearly laid out to volunteers before they agree to participate. The return for volunteers at all these scales is a key ingredient in the building of a critical mass of engaged citizens; not only for further experiments, but also to fulfill the objectives of the work that are not strictly scientific. Being a scientist, the direct relation with volunteers helps to improve the message and the way of delivering that message; to refine understanding of the phenomena involved in an experiment and even to refine a given experiment at future venues. One final positive side effect of this contact is the rise in public awareness of the difficulty and importance of science. As can be seen, forming the project around a rich and functional web page helps to harmonize all the time scales discussed and opens up new and interesting perspectives to bridge PUEs with other online-based CS practices. It also serves as an efficient way to communicate results and news of the project to interested citizens, and it could be used to improve data handling and sharing standards by allowing participants direct access to and management of their personal registers.

3. DISCUSSION

The advent of globalization and the fast track taken by innovation [48], combined with enormous challenges, have created demand for answers at a very fast pace. Deeply intertwined global and local actions are necessary to meet social challenges such as the continuous growth of the human population, the effects of climate change and even the need for collective decision-taking mechanisms prepared for effective policy-making. These urgent requirements collide with the typically long-winded process of scientific research, and this situation is affecting the philosophy, available resources and methods underlying science itself [55]. Society expects much from us as scientists but still lacks reflection and knowledge concerning the route toward more collaborative, public, open and responsive research. CS practices, even though they may not provide definitive answers to social challenges, aims to shorten the gap between the public and researchers, or in the worst case scenario, at least to increase social awareness of the problems tackled. CS practices can thus allow science to furnish ready-made solutions in public, and with citizens playing an active role.

In this work we have presented our experience in bridging the gap between CSS and the philosophy underlying CS, which in our case has taken the form of what we call PUEs. We hope that this work serves as a guideline for groups willing to adopt and expand such practices, and that it opens up the debate regarding the possibilities (and also the limitations) these approaches can offer. The flexibility of behavioral experiments can be combined with the strengths of big data to create a new tool capable of generating new collective knowledge. We have conceptually identified the main issues to be taken into account whenever CS research is planned in the field of CSS. We have grouped the challenges into three categories, in accordance with our experience: infrastructure, engagement and return. Furthermore, we have explained the solutions we implemented under the framework of a PUE, providing practical examples grounded in our own experience. The importance of team work and of widening the scope to consider questions that are not directly related with lab work has been highlighted, as well as the need to work hand in hand with both the public and other social actors. Other technical aspects of the approach related to necessary but peculiar infrastructure have also been reviewed.

We firmly support the idea of abandoning the *ivory tower* and opening up science and its research processes. Indeed, CS research essentially relies on the collaboration of citizens, but not just in passive data gathering. We think that placing distributed intelligence (with contributions from both experts and amateurs) at the very core of scientific analysis could also be a valid strategy to obtain rigorous and valuable results. We also believe that the PUEs we present here can potentially empower citizens to take their own civil action, relying on a collectively constructed facts-based approach [16]. To co-create and co-design a *smart city* with citizens, along the lines of the big data paradigm, will then be much easier and even more natural, as interests and concerns will be shared throughout the whole research process. Data gathered in the wild or *in vivo* contexts could thus be understood as truly public and open, while data ownership and knowledge would be shared from the very start [4]. Our future venues and experiments will be more deeply inspired by the open-source, do-it-yourself, do-it-together and makers movements [57], which facilitate learning-by-doing, and low-cost heuristic skills for everybody. A fresh look at problems can result in innovative and imaginative ideas that in the end can lead to *out of the box* solutions. However, as scientists, we will also have to find a way to reconcile unorthodox and intuitive forms with the standards and methodologies of the world of science. Lessons will need to be learnt from the “open prototyping” approach, in which an industrial product (such as a car) can be shaped by an iterative process during which the company owning the product has no problem allowing input from outside [58]. Some other clues can be found in the form of collective experimentation, where a fruitful dialogue can be established between the *matters of concern* raised by citizens and the *matters of fact* raised by scientists. Latour [29] already introduced these concepts and discussed their symbiotic relationship by considering the case study of ecologism (a civil movement) and ecology (a scientific activity). There are still many aspects to

test and explore concerning this approach in the field of CSS research.

We would also like, however, to present briefly open questions related to the way we perform science nowadays, and echo fundamental contradictions that science is not properly handling in this era of globalization [55]. CS practices yield pleasing outputs for communities; but they require a major effort from scientists, with the downside of providing very low (formal and bureaucratic) professional recognition. Open social experiments demand a high level of involvement in cooperation with non-scientific actors, which may divert professional researchers from the activity for which they are generally evaluated: the publication of results. Furthermore, such experiments often involve multidisciplinary teams, which then may encounter difficulties finding the appropriate journals to publish their findings and face difficulties with regard to acceptance in established communities. We thus urge the scientific community to actively recognize the valuable advantages of performing science within our proposed experimental framework.

Lessons learned must be shared, both within the community (as this paper attempts to do) and outside, in public spaces, including public institutions and policy makers. Science and CS are mostly publicly funded and therefore belong to society. The Internet provided new ways in which new relations among science and society can be strongly reinforced. We believe that this is good for everyone as it raises concern for science, by enhancing participation and, most importantly, by exploring new effective ways to push the boundaries of knowledge further. We hope that the present work helps in theoretically establishing the concept of the PUE and encouraging the adoption of CS practices in science, in whatever the field.

4. MATERIALS AND METHODS

In this section we provide descriptions of our experiences over 4 years of performing PUEs using CS practices in the city of Barcelona. Here we detail the experimental methods and briefly summarize the outcomes. All the experiments were performed in accordance with institutional (from the *Universitat de Barcelona* in all cases except for the Cooperation(1) experiment, when they were from the *Universidad Carlos III*) and national guidelines and regulations concerning data privacy (in accordance with Spanish data protection law: the LOPD) and gave written informed consent in accordance with the Declaration of Helsinki. All interfaces used included informed consent from all subjects. The data collected were properly anonymized and not related to personal details, which in our case were age range, sex, level of education and electronic (e-mail) address.

4.1. The Bee-Path(1) Experiment

The aim of this experiment was to study the movement of visitors during their exploration of an outdoor science and technology fair where several stands with activities were located in an area of approximately 3 h inside a public park. The experiment took place during the weekend of 16th and 17th June 2012, specifically on Saturday afternoon (from 16 to 20 h) and the morning of Sunday (from 11 to 15 h). The participants had very different

interests, origins, backgrounds and ages; the organization of the event estimated that 10,000 people visited the fair. The Bee-Path information stand was located at the main entrance where visitors were encouraged to participate in the experiment by downloading an App onto their mobile phones. After very simple registration and instructions on how to activate the App, the participants were left to wander around the fair while being tracked. After a laborious cleaning process, we analyzed the movement and trajectories of 27 subjects from the records provided by the 101 volunteers. We found spatio-temporal patterns in their movement and we developed a theoretical model based on Langevin dynamics driven by a gravitational potential landscape created by the stands. This model was capable of explaining the results of the experiment and predicted scenarios with other spatial configurations of the stands. A scientific paper has been submitted for publication [59]. The project description, results and data are freely accessible on the web page: www.bee-path.net.

4.2. The Bee-Path(2) Experiment

Following the previous Bee-Path(1) experiment, this set-up also focussed on studying the movement of people, but in this case we were interested in searching patterns or how people move when exploring a landscape to find something. The experiment took place at the following edition of the same science and technology fair (June 15th and 16th 2013). The participants also downloaded an App, similar to that used in Bee-Path(1), that tracked them; but in this case they were instructed to find 10 dummies hidden in the park. Numerous problems were encountered due to several technological and non-technological factors which impeded satisfactory performance of the experiment. The technological issues were principally two: the low accuracy of the recordings, due to the proximity of the regional parliament building where Wi-Fi and mobile phone coverage was inhibited; and the limited performance of the App when running in low-end devices. Non-technological problems included a bad placement of our stand (far from the entrance) where recruitment of volunteers was hard; and the unexpected fact that some members of the public altered, stole or changed the positions of some of the dummies. Notwithstanding the failure of the experiment to produce meaningful and useful results, we learnt important lessons from these complications.

4.3. The Cooperation(1) Experiment

Here we explored how important age is in the emergence of cooperation when people repeatedly face the *prisoner's dilemma* (PD). The experiment was carried out with 168 volunteers selected from the attendants of the Barcelona DAU Festival 2012 (Barcelona's 1st board game fair; December 15th and 16th). The set of volunteers was divided into 42 subsets of 4 players according to age: seven different age groups plus one control group in which the subjects were not distinguished by age. Each subset took part in a game where the four participants played 25 rounds (although they were not aware of it) deciding between two colors associated with a certain PD pay-off matrix. The participants played a 2×2 PD game with each of their 3 neighbors, choosing the same action for all opponents. In

order to play with an incentive, they were remunerated with real money proportionally to their final score. During the game, volunteers interacted through software specially programmed for the experiment and installed on a laptop. They were not allowed to talk or signal in any way, but to further guarantee that potential interactions among the players would not influence the results of the experiment, the assignment of players to the different computers in the room was completely random. In this experiment, together with the "Jesuites Casp" experiment described in the next subsection, we found that the elderly cooperated more, and there is a behavioral transition from reciprocal but more volatile behavior, to more persistent actions toward the end of adolescence. For further details see [60].

4.4. The Cooperation(2) Experiment

The purpose of repeating the Cooperation(1) experiment at the DAU festival was to confirm the apparent tendency of children to cooperate less than the average population. Thus, we repeated the experiment simply to increase the pool of subjects in this age range, which allowed us to be more statistically accurate. We analyzed the performance of 52 secondary school children (the "Jesuites Casp" experiment) ranging from 12 to 13 years old. The methods and protocols were the same as in the Cooperation(1) experiment, as was the software installed on the laptops. The results of this experiment refuted the hypothesis that children cooperate less on average, but at the same time confirmed their more volatile behavior, as described in Gutiérrez-Roig et al. [60].

4.5. The Mr. Banks(1) Experiment

The Mr. Banks experiment was set up to study how non-expert people make decisions in uncertain environments; specifically, assessing their performance when trying to guess if a real financial market price will go up or down. We analyzed the performance of 283 volunteers at the Barcelona DAU Festival 2013 (from approximately 6000 attendants) on December 14th and 15th. All the volunteers played via an interface that was specifically created for the experiment and was accessible via identical tablets only available in a specific room under researcher surveillance. On the main screen the devices showed the historic daily market price curve and some other information such as 5-day and 30-day average window curves, the high-frequency price on the previous day, the opinion of an expert, the price direction on previous days and price directions of other markets around the world. All the price curves and information were extracted from real historical series. The participants could play in four different scenarios with different time and information availability constraints. In each scenario they were required to make guesses for 25 rounds, while every click on the screen was recorded. Each player started with 1000 coins and earned an additional 5% of their current score if their guess was correct or lost an equivalent negative return if their guess was wrong. We used gamification strategies and we did not provide any economic incentive, in contrast to the social dilemma experiments: Cooperation(1), Cooperation(2), and Dr. Brain. The analysis of the 18,436 recorded decisions and 44,703 clicks allowed us to conclude that participants tend to follow intuitive strategies called "market imitation" and "win-stay, lose-switch." These strategies are followed less closely when

there is more time to make a decision or some information will be provided (Gutiérrez-Roig et al., submitted). Both the experiment and information on the project are available at: www.mr-banks.net.

4.6. The Mr. Banks(2) Experiment

Here we repeated the Mr. Banks(1) experiment in a different context, in order to study the reproducibility of the results. The interface and the experimental set-up were the same as in Mr. Banks(1) but the typology of participants and the type of event were significantly different. The experiment, named Hack your Brain in the conference programme, was situated at the main entrance of the CAPS2015 conference: the International Event on Collective Awareness Platforms for Sustainability and Social Innovation, held in Brussels (July 7th and 8th 2015). The 42 volunteers who played the game provided 2372 recorded decisions. The volunteers were all registered participants at the conference with very diverse profiles (scientists, mostly from the social sciences; social innovators; designers; social entrepreneurs; policymakers; etc.). The results showed that the results of the Mr. Banks(1) experiment had a good reproducibility, as the percentage of correct guesses was similar in Mr. Banks(1) and Mr. Banks(2) (53.4 and 52.7%, respectively) as was the percentage of market-up decisions (60.8 and 60.5%, respectively). Deeper analysis of the results is underway, to check that the strategies adopted by Mr. Banks(1) volunteers were the same as those used by Mr. Banks(2) volunteers. A paper has been submitted for publication: (Gutiérrez-Roig et al., submitted).

4.7. The Dr. Brain Experiment

This was a lab-in-the-field experiment that allows for a phenotypic characterization of individuals when facing different social dilemmas. Instead of playing with the same fixed pay-off matrix, as in the Cooperation experiments, here the values and the neighbors changed every round. We discretized the (T, S) -plane as a lattice of 11×11 sites, allowing us to explore up to 121 different games grouped in 4 categories: Harmony Games, Stag Hunt Games, Snowdrift Games and Prisoner's Dilemma Games. Each player was given a tablet with the App for the experiment installed. The participants were shown a brief tutorial, but were not instructed in any particular way, nor with any particular goal in mind. They were informed that they had to make decisions under different conditions and against different opponents in every round. Due to practical limitations, we could only host around 25 players simultaneously, so the experiment was conducted in several sessions over a period of 2 days. In every session, all the individuals played a different number of rounds, picked at randomly between 13 and 18. The total number of participants in our experiment was 541, and a total of 8366 game actions were collected. In order to play with an incentive, they received coupons for a prize of EUR 50 to spend in neighborhood shops. During the game, the volunteers interacted through software specially programmed for the experiment and

installed in a laptop. They were not allowed to talk or signal in any other way, and again were spatially placed at random. From this experiment we concluded that we can distinguish, empirically and without making any assumptions, five different types of player's behavior or phenotypes that are not theoretically predicted. A paper has been submitted for publication: (Poncela-Casasnovas et al., submitted).

AUTHOR CONTRIBUTIONS

OS, MG, IB, and JP equally conceived and wrote the work. All the authors approved the final version of the manuscript.

FUNDING

The research leading to these results received funding from: Barcelona City Council (Spain); RecerCaixa (Spain) through grant *Citizen Science: Research and Education*; MINECO (Spain) through grants FIS2013-47532-C3-2-P and FIS2012-38266-C2-2; Generalitat de Catalunya (Spain) through grants 2014-SGR-608 and 2012-ACDC-00066; and Fundación Española para la Ciencia y la Tecnología (FECYT, Spain) through the Barcelona Citizen Science Office project of the Barcelona Lab programme. OS also acknowledges financial support from the Generalitat de Catalunya (FI programme) and the Spanish MINECO (FPU programme).

ACKNOWLEDGMENTS

We would like to acknowledge the participation of at least 1255 anonymous volunteers who made this research possible. We especially thank Mar Canet, Nadala Fernández, Oscar Marín from Outliers, Carlota Segura, Clàudia Payrató, Pedro Lorente, Fran Iglesias, David Roldán, Marc Anglès, and Berta Paco for all the logistics and their full collaboration which made the experiments possible in one way or another. We also acknowledge the co-authors Anxo Sánchez, Yamir Moreno, Carlos Gracia-Lázaro, Jordi Duch, Julián Vicens, Julia Poncela-Casasnovas, Jesús Gómez-Gardenes, Albert Díaz-Guilera, Federic Bartomeus, Aitana Oltra, and John Palmer in the research papers following the experiments reported here. We also thank the director of the DAU (Oriol Comas) for giving us the opportunity to perform the three experiments at the DAU Barcelona Festival. We are greatly indebted to the Barcelona Lab programme, promoted by the Direction of Creativity and Innovation of the Barcelona City Council led by Inés Garriga for their help and support in setting up the experiments at the Barcelona Science Fair (Parc de la Ciutadella, public park in Barcelona) and at the DAU Barcelona Festival (Fabra i Coats, Creativity Fabrique of the City Council). Finally, we also thank Marta Arniani from Sigma-Orionis for hosting the Mr. Banks(2) experiment at the CAPS event.

REFERENCES

1. Latour B. From the world of science to the world of research? *Science* (1998) **280**:208–9. doi: 10.1126/science.280.5361.208
2. Kuhn TS. *The structure of scientific revolutions*. Chicago, IL: University of Chicago Press (2012).
3. Latour B. Which protocol for the new collective experiments? (2006) 32–3. Available online at: <http://www.bruno-latour.fr/sites/default/files/P-95-METHODS-EXPERIMENTS.pdf>
4. Callon M, Rabeharisoa V. Research “in the wild” and the shaping of new social identities. *Technol Soc.* (2003) **25**:193–204. doi: 10.1016/S0160-791X(03)00021-6
5. Hand E. People power. *Nature* (2010) **466**:685–7. doi: 10.1038/466685a
6. Bonney R, Shirk JL, Phillips TB, Wiggins A, Ballard HL, Miller-Rushing AJ, et al. Next steps for citizen science. *Science* (2014) **343**:1436–7. doi: 10.1126/science.1251554
7. Gura T. Citizen science: amateur experts. *Nature* (2013) **496**:259–61. doi: 10.1038/nj7444-259a
8. Editorial. Rise of the citizen scientist. *Nature* (2015) **524**:265. doi: 10.1038/524265a
9. Societize. *White Paper on Citizen Science, Societize*. Available online at: <http://ec.europa.eu/digital-agenda/en/news/green-paper-citizen-science-europe-towards-society-empowered-citizens-and-enhanced-research-0;2014>
10. Irwin A. *Citizen science: A Study of People, Expertise and Sustainable Development*. New York, NY: Psychology Press (1995).
11. Lintott CJ, Schawinski K, Slosar A, Land K, Bamford S, Thomas D, et al. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *Mon Not R Astron Soc.* (2008) **389**:1179–89. doi: 10.1111/j.1365-2966.2008.13689.x
12. Kim JS, Greene MJ, Zlateski A, Lee K, Richardson M, Turaga SC, et al. Spacetime wiring specificity supports direction selectivity in the retina. *Nature* (2014) **509**:331–6. doi: 10.1038/nature13240
13. Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, et al. Predicting protein structures with a multiplayer online game. *Nature* (2010) **466**:756–60. doi: 10.1038/nature09304
14. Lauro FM, Sensi SJ, Cullen J, Neches R, Jensen RM, Brown MV, et al. The Common oceanographer: crowdsourcing the collection of oceanographic data. *PLoS Biol.* (2014) **12**:e1001947. doi: 10.1371/journal.pbio.1001947
15. Haklay M. Citizen science and volunteered geographic information: overview and typology of participation. In: Sui D, Elwood S, Goodchild M, editors. *Crowdsourcing Geographic Knowledge*. Dordrecht: Springer (2013). pp. 105–122.
16. McQuillan D. The countercultural potential of citizen science. *M/C J.* (2014) **17**. Available online at: <http://journal.media-culture.org.au/index.php/mcjournal/article/viewArticle/919>
17. Lazer D, Pentland A, Adamic L, Aral S, Barabasi AL, Brewer D, et al. Social science. Computational social science. *Science* (2009) **323**:721–3. doi: 10.1126/science.1167742
18. Cioffi-Revilla C. Computational social science. *Wiley Interdiscip Rev.* (2010) **2**:259–71. doi: 10.1002/wics.95
19. Russell SA. *Diary of a Citizen Scientist: Chasing Tiger Beetles and Other New Ways of Engaging the World*. Corvallis, OR: Oregon State University Press (2014).
20. Dickinson JL, Bonney R. *Citizen Science: Public Participation in Environmental Research*. New York, NY: Comstock Publishing Associates (2015).
21. Silvertown J. A new dawn for citizen science. *Trends Ecol Evol.* (2009) **24**:467–71. doi: 10.1016/j.tree.2009.03.017
22. Hoggatt AC. An experimental business game. *Behav Sci.* (1959) **4**:192–203. doi: 10.1002/bs.3830040303
23. Cueva C, Roberts RE, Spencer T, Rani N, Tempest M, Tobler PN, et al. Cortisol and testosterone increase financial risk taking and may destabilize markets. *Sci Rep.* (2015) **5**:11206. doi: 10.1038/srep11206
24. Yoshimura Y, Sobolevsky S, Ratti C, Girardin F, Carrascal JP, Blat J, et al. An analysis of visitors behavior in the louvre museum: a study using bluetooth data. *Environ Plan B* (2014) **41**:1113–31. doi: 10.1068/b130047p
25. Grujić J, Fosco C, Araujo L, Cuesta JA, Sánchez A. Social experiments in the mesoscale: humans playing a spatial prisoner’s dilemma. *PLoS ONE* (2010) **5**:e13749. doi: 10.1371/journal.pone.0013749
26. Weiss CH, Poncela-Casasnovas J, Glaser JI, Pah AR, Persell SD, Baker DW, et al. Adoption of a high-impact innovation in a homogeneous population. *Phys Rev X* (2014) **4**:041008. doi: 10.1103/PhysRevX.4.041008
27. Kagel JH. *The Handbook of Experimental Economics*. Princeton, NJ: Princeton University Press (1997).
28. Margolin D, Lin YR, Brewer D, Lazer D. Matching data and interpretation: towards a rosetta stone joining behavioral and survey data. In: *Seventh International AAAI Conference on Weblogs and Social Media*. Boston, MA (2013).
29. Latour B. *Politics of Nature*. Cambridge, MA: Harvard University Press (2009).
30. Gracia-Lázaro C, Ferrera A, Ruiza G, Tarancón A, Cuesta JA, Sánchez A, et al. Heterogeneous networks do not promote cooperation when humans play a Prisoner’s Dilemma. *Proc Natl Acad Sci USA.* (2012) **109**:12922–26. doi: 10.1073/pnas.1206681109
31. Perelló J, Murray-Rust D, Nowak A, Bishop SR. Linking science and arts: intimate science, shared spaces and living experiments. *Eur Phys J Spec Top.* (2012) **214**:597–634. doi: 10.1140/epjst/e2012-01707-y
32. Cattuto C, Van den Broeck W, Barrat A, Colizza V, Pinton JF, Vespignani A. Dynamics of person-to-person interactions from distributed RFID sensor networks. *PLoS ONE* (2010) **5**:e11596. doi: 10.1371/journal.pone.0011596
33. Schroeder R. Big Data and the brave new world of social media research. *Big Data Soc.* (2014) **1**:2053951714563194. doi: 10.1177/2053951714563194
34. Conte R, Gilbert N, Bonelli G, Cioffi-Revilla C, Deffuant G, Kertesz J, et al. Manifesto of computational social science. *Eur Phys J Spec Top.* (2012) **214**:325–46. doi: 10.1140/epjst/e2012-01697-8
35. Lazer D, Kennedy R, King G, Vespignani A. The parable of google flu: traps in big data analysis. *Science* (2014) **343**:1203–5. doi: 10.1126/science.1248506
36. Boyd D, Crawford K. Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. *Inf Commun Soc.* (2012) **15**:662–79. doi: 10.1080/1369118X.2012.678878
37. Vianna GMS, Meekan MG, Bornovski TH, Meeuwig JJ. Acoustic telemetry validates a citizen science approach for monitoring sharks on coral reefs. *PLoS ONE* (2014) **9**:e95565. doi: 10.1371/journal.pone.0095565
38. Rossiter DG, Liu J, Carlisle S, Zhu AX. Can citizen science assist digital soil mapping? *Geoderma* (2015) **259–260**:71–80. doi: 10.1016/j.geoderma.2015.05.006
39. Bowser A, Wiggins A, Shanley L, Preece J, Henderson S. Sharing data while protecting privacy in citizen science. *Interactions* (2014) **21**:70–73. doi: 10.1145/2540032
40. Resnik DB, Elliott KC, Miller AK. A framework for addressing ethical issues in citizen science. *Environ Sci Policy* (2015) **54**:475–81. doi: 10.1016/j.envsci.2015.05.008
41. Riesch H, Potter C. Citizen science as seen by scientists: methodological, epistemological and ethical dimensions. *Public Underst Sci.* (2013) **23**:107–20. doi: 10.1177/0963662513497324
42. Prestopnik NR, Crowston K. Gaming for (Citizen) science. In: *Seventh IEEE International Conference on e-Science Workshops*. New York, NY (2011). pp. 28–33.
43. Sauermann H, Franzoni C. Crowd science user contribution patterns and their implications. *Proc Natl Acad Sci USA.* (2015) **112**:679–84. doi: 10.1073/pnas.1408907112
44. Curtis V. *Online Citizen Science Projects: An Exploration of Motivation, Contribution and Participation*. Ph.D. thesis. The Open University (2015).
45. Petersen JC. *Citizen Participation in Science Policy*. Amherst, MA: University of Massachusetts Press (1984).
46. Elster J. *Deliberative Democracy*. Cambridge, MA: Cambridge University Press (1998).
47. *EU Commission. Responsible Research and Innovation, Europeans Ability to Respond to Societal Challenges* (2014). (Accessed: 2015-07-29). Available online at: http://ec.europa.eu/research/science-society/document_library/pdf_06/responsible-research-and-innovation-leaflet_en.pdf
48. Chesbrough HW. *Open innovation: The New Imperative for Creating and Profiting From Technology*. Cambridge, MA: Harvard Business Press (2006).
49. Henrich J, Heine SJ, Norenzayan A. Most people are not WEIRD. *Nature* (2010) **466**:29. doi: 10.1038/466029a

50. Collaboration OS. Estimating the reproducibility of psychological science. *Science* (2015) **349**:aac4716. doi: 10.1126/science.aac4716
51. Bohannon J. Many psychology papers fail replication test. *Science* (2015) **349**:910–1. doi: 10.1126/science.349.6251.910
52. Webster MJ, Sell J. *Laboratory Experiments in the Social Sciences*. Boston, MA: Academic Press (2014).
53. Kramer ADI, Guillory JE, Hancock JT. Experimental evidence of massive-scale emotional contagion through social networks. *Proc Natl Acad Sci USA*. (2014) **111**:8788–90. doi: 10.1073/pnas.1320040111
54. Gleibs IH. Turning virtual public spaces into laboratories: thoughts on conducting online field studies using social network sites. *Anal Soc Issues Publ Policy* (2014) **14**:352–70. doi: 10.1111/asap.12036
55. Franzoni C, Saueremann H. Crowd science: the organisation of scientific research in open collaborative projects. *Res Policy* (2014) **43**:1–20. doi: 10.1016/j.respol.2013.07.005
56. Morris B, Croker S, Zimmerman C, Gill D, Romig C. Gaming science: the ‘Gamification’ of scientific thinking. *Front Psychol*. (2013) **4**:607. doi: 10.3389/fpsyg.2013.00607
57. Hatch M. *The Maker Movement Manifesto*. New York, NY: McGraw-Hill Education (2014).
58. Bullinger AC, Hoffmann H, Leimeister JM. The next step-open prototyping. In: *ECIS*. Atlanta, GA (2011).
59. Gutiérrez-Roig M, Sagarra O, Oltra A, Palmer JRB, Bartumeus F, Diaz-Guilera A, et al. Active and reactive behaviour in human mobility: the influence of attraction points on pedestrians. ArXiv: 1511.03604 (2015).
60. Gutiérrez-Roig M, Gracia-Lázaro C, Perelló J, Moreno Y, Sánchez A. Transition from reciprocal cooperation to persistent behaviour in social dilemmas at the end of adolescence. *Nat Commun*. (2014) **5**:4362. doi: 10.1038/ncomms5362

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Sagarra, Gutiérrez-Roig, Bonhoure and Perelló. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Mathematical modeling of complex contagion on clustered networks

David J. P. O'Sullivan*, Gary J. O'Keeffe, Peter G. Fennell and James P. Gleeson

Mathematics Applications Consortium for Science and Industry, Department of Mathematics and Statistics, University of Limerick, Limerick, Ireland

OPEN ACCESS

Edited by:

Javier Borge-Holthoefer,
Qatar Computing Research Institute,
Qatar

Reviewed by:

Nicola Perra,
University of Greenwich, UK
Samuel Johnson,
University of Warwick, UK

*Correspondence:

David J. P. O'Sullivan,
Department of Mathematics and
Statistics, University of
Limerick, A2016g, Limerick, Ireland
david.osullivan@ul.ie

Specialty section:

This article was submitted to
Interdisciplinary Physics,
a section of the journal
Frontiers in Physics

Received: 05 July 2015

Accepted: 21 August 2015

Published: 15 September 2015

Citation:

O'Sullivan DJP, O'Keeffe GJ,
Fennell PG and Gleeson JP (2015)
Mathematical modeling of complex
contagion on clustered networks.
Front. Phys. 3:71.
doi: 10.3389/fphy.2015.00071

The spreading of behavior, such as the adoption of a new innovation, is influenced by the structure of social networks that interconnect the population. In the experiments of Centola [15], adoption of new behavior was shown to spread further and faster across clustered-lattice networks than across corresponding random networks. This implies that the “complex contagion” effects of social reinforcement are important in such diffusion, in contrast to “simple” contagion models of disease-spread which predict that epidemics would grow more efficiently on random networks than on clustered networks. To accurately model complex contagion on clustered networks remains a challenge because the usual assumptions (e.g., of mean-field theory) regarding tree-like networks are invalidated by the presence of triangles in the network; the triangles are, however, crucial to the social reinforcement mechanism, which posits an increased probability of a person adopting behavior that has been adopted by two or more neighbors. In this paper we modify the analytical approach that was introduced by Hébert-Dufresne et al. [19], to study disease-spread on clustered networks. We show how the approximation method can be adapted to a complex contagion model, and confirm the accuracy of the method with numerical simulations. The analytical results of the model enable us to quantify the level of social reinforcement that is required to observe—as in Centola’s experiments—faster diffusion on clustered topologies than on random networks.

Keywords: clustered networks, complex contagion, clique networks, clique approximation, social reinforcement, diffusion of behavior

1. Introduction

Many systems find a natural interpretation as a complex network where nodes identify the objects of the system and the links between nodes represent the presence of a relationship or interaction between those objects [1]. Such network characterizations range from friendships on Facebook [2], connections between web-pages by hyper-links [3], to protein interaction networks in biological systems [4]. A growing area of interest is the modeling of how behaviors diffuse across social networks, such as the adoption of innovations [5] or the spreading of information [6, 7]. Epidemiological models provide a convenient architecture for articulating these spreading processes where nodes (individuals) can be in one of two states adopter (“infected”) or non-adopter (“susceptible”).

The diffusion of social behavior is often characterized as either a “simple contagion” or a “complex contagion” [8]. A simple contagion is any process where a node can easily become infected by a single contact with an infected neighbor; on the other hand a complex contagion is a process where a node usually requires multiple exposures before they change state [9]. Simple

contagions arise naturally in disease spread-models where a susceptible individual only requires a single contact with an infected individual to allow a pathogen to propagate. Traditionally simple contagion models have been applied to sociological spreading behaviors in order to predict how a behavior would diffuse across a network [10]. The simplest example is the SI (susceptible-infected) model, for example, where infected nodes transmit infection across their links at a rate β per unit time [11]. Susceptible nodes change state and become infected (i.e., adopt the behavior) at a rate that scales linearly with the number of infected network neighbors (see Section 3 for more details). Once infected, a node cannot recover to the susceptible state (an adopter node can not unadopt a behavior); the SI model therefore provides an example of a binary-state monotone dynamic process [11].

The importance of network topology for spreading dynamics, specifically the density of triangles (clustering) in the network, has been well established [12]. In social networks, clustering provides a useful measure for how densely connected local groups are [13]. A high density of triangles implies a high chance that “the friend of my friend is also a friend of mine.” It has been shown that the lower the density of triangles the further a simple contagion will spread across a network [14], because each additional infected node has a high chance of linking to unexposed nodes. Conversely, a high density of triangles results in a slower spread because the disease travels across “redundant” links to nodes that have already been infected [15]. The ideal case for efficient propagation of a simple contagion is a random network where each node’s links connect to different neighborhoods; random networks necessarily have no presence of clustering in the topology. If a simple contagion model (such as the SI model) accurately describes the spreading of social behaviors then we should observe faster diffusion of such behaviors on networks with lower clustering. However, in a groundbreaking experiment by Centola [15], the opposite was observed. Centola found that the diffusion of adoption spread further and faster on networks with a high degree of clustering than on corresponding (same mean degree) random networks, contradicting the results predicted by simple contagion models. He observed that nodes who received multiple exposures to the behavior were more likely to adopt than those who had only received one exposure, indicating that the behavior spread as a complex contagion.

In this paper we present a complex contagion model that reflects the requirement for multiple exposures to effectively propagate a behavior through a clustered network. Using the complex contagion model we examine the spreading behavior produced on networks with varying levels of clustering. Lü et al. [16] have also numerically examined models for adoption, but only on small networks, whereas we concentrate on the large-network limit ($N \rightarrow \infty$) where analytic results can be found. Modeling simple contagions on random networks is well understood, where analytic results for the fraction of infected nodes in the steady state are relatively easy to calculate by standard approximation schemes such as mean-field (MF) or pair-approximation (PA) methods [17]. However, accurately approximating diffusion processes on clustered

networks remains a challenge. The presence of clustering immediately invalidates the assumption of locally tree-like network structure that MF and PA methods are based upon [18]. In our context, the presence of triangles is integral to the reinforcement mechanism of a complex contagion. To address this we modify the analytic approach introduced by Hébert-Dufresne et al. [19]. Their framework was used to model disease-spread processes on clustered networks. We show how the approximation method can be adapted to a complex contagion model, and confirm the accuracy of the method with numerical simulations.

The remainder of the paper is structured as follows. The clique-based network that forms the basis for our examinations of complex contagion is outlined in Section 2. The complex contagion model is described in Section 3. Section 4 presents the approximation scheme that is used to account for presence of clustering and the procedure for finding a linearized solution to the system. In Section 5 we examine the accuracy of the approximation and the results of the complex contagion model. Finally, Section 6 presents our conclusions.

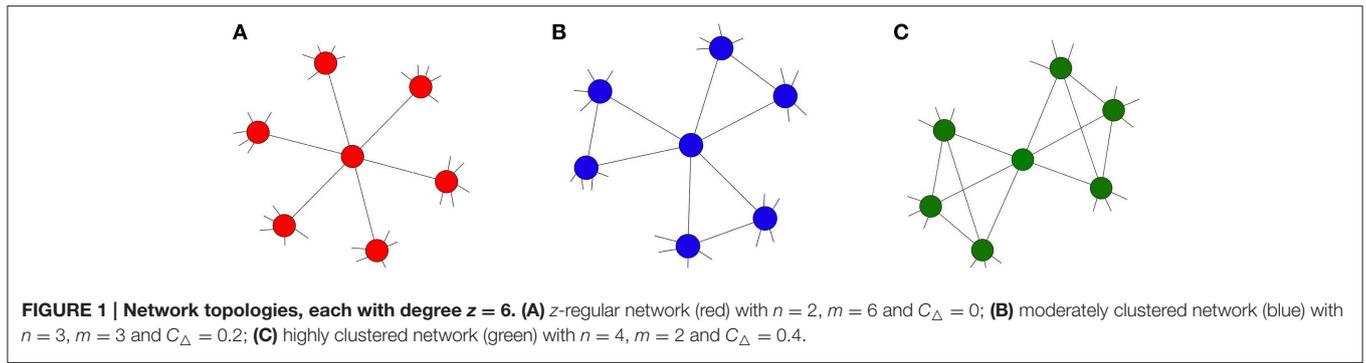
2. Clique-based Networks

The defining characteristic of a complex contagion is the increased propensity to become infected (adopt) a behavior given multiple exposures [9]. We expect to observe different spreading behavior of a complex contagion depending on the level of clustering on the network. This is because there is a higher propensity on clustered networks for multiple infected nodes to have a susceptible node in common when compared against random networks. Therefore, clustering is the salient feature of a network that we wish to isolate. To quantify the clustering in a network we use the global clustering coefficient [20], defined as

$$C_{\Delta} = \frac{3 \times N_{\Delta}}{N_3}, \quad (1)$$

where N_{Δ} is the total number of triangles in the network and N_3 is the number of connected triples of nodes. The case $C_{\Delta} = 0$ implies that no paths of length three are closed, meaning that the network is locally tree-like [21].

When examining the diffusion produced on differing networks we must be careful to compare like with like so as not to introduce confounding factors into our analysis. Therefore, we use networks that allow us to control the clustering, while holding other topological features (such as the degree distribution) constant; this is achieved using clique-based networks [1, 19, 22]. In a clique-based network, each clique has n (randomly-chosen) nodes and each node is a part of m (randomly-chosen) cliques. For example, a triangle is a clique with $n = 3$ nodes. Use of these networks follows, in spirit, the experimental design used by Centola, where clustered lattices were compared to z -regular random networks of the same degree to isolate the effects of clustering (see Appendix A for details on his experiment). However, the clique-based networks allow us to use analytical methods that cannot be directly applied to the clustered lattice networks used by Centola.

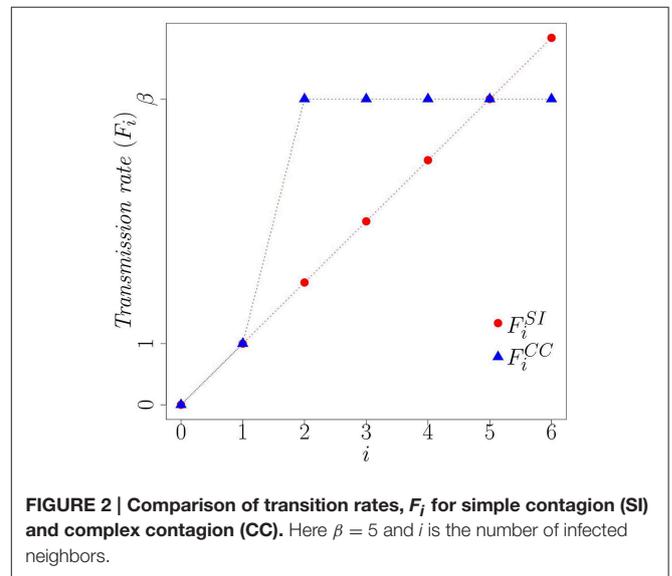


We examine various different forms of clique motifs. This is done by varying n and m subject to the constraint that the degree of each node is fixed, specifically, the degree of each node is $z = (n - 1)(m) = 6$ (as in Centola's main experiments). We focus on three motif types which are illustrated in **Figure 1**. The motif in **Figure 1A** corresponds to a random network where each clique contains two nodes and each node is part of six cliques ($n = 2$ and $m = 6$), i.e., each "clique" is just a link in the random 6-regular network. The motif shown in **Figure 1B** is a triangle, with each node being part of three cliques ($n = 3$ and $m = 3$). The last motif in **Figure 1C** is a four clique and each node is part of two cliques ($n = 4$ and $m = 2$). These local topologies result in networks with clustering coefficients of 0, 0.2, and 0.4, respectively. As each network is constructed from the aforementioned motifs, there is no variation in degree or local clustering between nodes. Thus, we can isolate the effect of clustering on the spread of a complex contagion between the different networks. In the next section we define our complex contagion model which will capture the defining characteristic of a complex contagion where nodes that receive multiple exposures have an increased propensity to change state over those who have received only one.

3. Complex Contagion Model

In this section we define our complex contagion model. First we briefly define the susceptible-infected (SI) model for comparison purposes. In the case of the continuous-time SI model (which is a simple contagion model) an infected node transmits disease to all its network neighbors at a rate β , where a neighboring node's probability of changing state from this contact is βdt in an infinitesimal time interval of length dt . A susceptible node with i infected neighbors therefore is exposed to i independent sources of infection, so the probability that the node does not become infected in a time interval dt is $(1 - \beta dt)^i$, with the probability that the node does become infected being $1 - (1 - \beta dt)^i$. We define the *transition rate* F_i^{SI} by letting the probability of infection in a small-time interval dt equal $F_i^{SI} dt$. As $dt \rightarrow 0$ this probability becomes, $\beta i dt$, and so the transition rate for a node with i infected neighbors is

$$F_i^{SI} = \beta i. \tag{2}$$



The transition rate scales linearly with the number i of infected neighbors, which is reasonable for a biological contagion where each possible infection event is independent of the others. However, for a social contagion a node will rarely adopt a behavior after a single exposure, it is only after several exposures that a node becomes likely to adopt [15].

As a deliberately simplified model for complex contagion we therefore propose the following transition rate function:

$$F_i^{CC} = \begin{cases} 0 & \text{if } i = 0, \\ 1 & \text{if } i = 1, \\ \beta & \text{if } i > 1. \end{cases} \tag{3}$$

where β is the rate at which a susceptible node changes state, given multiple exposures. To model complex contagion with strong social reinforcement, for example, we can set $\beta \gg 1$. **Figure 2** compares the transition rates for SI and Complex Contagion [Equations (2) and (3), respectively] as a function of the number i of infected neighbors of a susceptible node. Considering Equation (3) and assuming $\beta \geq 2$, if a node has multiple infected neighbors ($i \geq 2$) it has an increased propensity to adopt in comparison to a node with only one

infected neighbor ($i = 1$). In an experiment where the contagion begins with a small fraction of infected nodes the chance that a node will receive multiple exposures is much higher on a clustered network than on a random network, resulting in faster spread over clustered topologies. As we show below, this very simple representation of a complex contagion can capture the spreading behavior observed by Centola while still remaining amenable to mathematical analysis.

4. Clique Approximation

4.1. Clique Approximation Scheme

Many approximation schemes have been developed in order to help approximate the relationships between macroscopic observables (such as the fraction of nodes infected) and stochastic microscopic (node-level) events, such as the number of infected neighbors of each node. Such approximation schemes vary in their level of complexity, with an inherent trade-off between accuracy and complexity. There are two main approximation schemes, the *mean-field* (MF) and *pair-approximation* (PA) methods.

Briefly, the MF approximation assumes that the states of every node in the network are independent. Pair-approximation (PA) methods extend the MF approximation to incorporate information about the pair-wise correlations between susceptible nodes and their neighbors' states. For a more detailed discussion of these methods refer to Porter and Gleeson [11] and references therein. The MF and PA methods assume that the networks are locally tree-like (absence of local clustering). Violations of this assumption results in poor approximations to the true behavior of the spreading dynamics. As clustering is an integral part of the networks we consider here, we require the development of an analytical framework that can take into account both the complex contagion and the presence of clustering in clique-type networks. We will refer to this as the clique approximation (CA) scheme. **Figure 3** provides a schematic of the level of local topology that each approximation scheme takes into account.

We extend the method introduced by Hebert-Dufresne et al. [19] which they used to study SIS disease-spread dynamics (where an infected node can transition back to the susceptible state) on clique-styled networks. Our initial focus is on extending their method from simple contagions to apply it to complex

contagion models such as Equation (3). In the CA scheme we track the time-dependent fraction $c_i(t)$ of cliques that contain i infected nodes, where the transition of a clique with i infected nodes to a clique with $i + 1$ infected nodes is described by the time-dependent transition rate $\gamma_i(t)$, as illustrated in **Figure 4**. Recall from Section 2 that the networks we examine are created from basic motifs where each clique had n nodes and each node is part of m cliques. Consequently, the networks are (1) z -regular (all nodes have the same degree) and (2) each node has the same local topology (refer to **Figure 1** for examples).

Tracking the dynamical states of cliques, as opposed to nodes, results in a more complicated system of equations than the MF or PA methods. The added complexity is required to account for the presence of clustering in the network. We wish to calculate the fraction of infected nodes at time t , which we denote $\rho(t)$. To create an evolution equation for $\rho(t)$ we first calculate the rate of change of the fraction $c_i(t)$ of cliques with i infected nodes at time t . Note the normalization condition $\sum_{i=0}^n c_i = 1$ applies at all times t . The number of nodes that can leave a clique in state c_{i-1} and enter state c_i is the total number n of nodes in that clique minus the number of nodes that are already infected at time t , i.e., $n - (i - 1)$. Similarly, the fraction of nodes that can leave a clique in state c_i and move to a clique in state c_{i+1} is $(n - i)c_i$. Applying the relevant transition rates (γ_{i-1} and γ_i , respectively) at which nodes change from one clique class to another (**Figure 4**) results in:

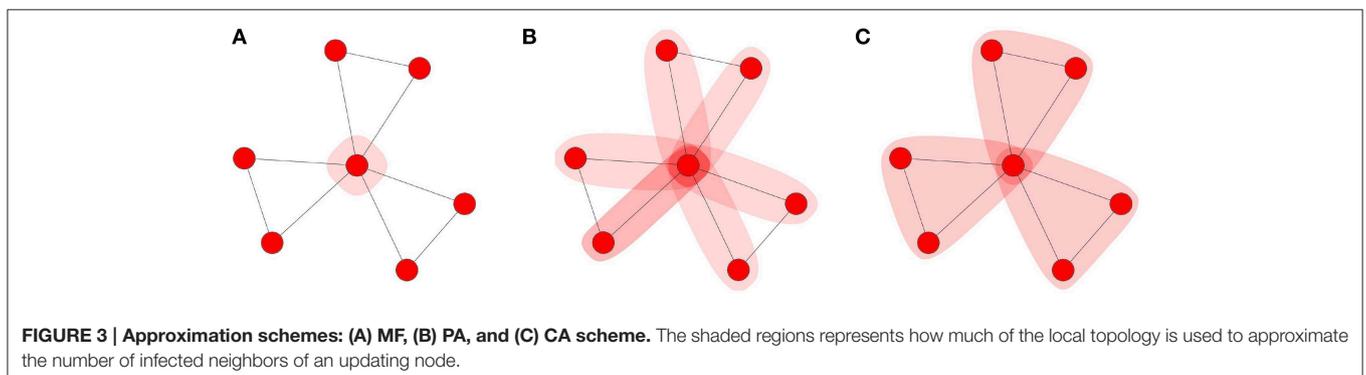
$$\begin{aligned} \frac{dc_i(t)}{dt} = & (n - i + 1)c_{i-1}(t)\gamma_{i-1}(t) \\ & - (n - i)c_i(t)\gamma_i(t) \end{aligned} \tag{4}$$

for $i = 0, 1, \dots, n$.

(Note the explicit dependence of variables on t is henceforth omitted for convenience.) Using Equation (4) we can calculate $d\rho/dt$ by realizing that each clique with i infected nodes contributes i/n nodes to the total fraction of infected nodes:

$$\frac{d\rho}{dt} = \frac{1}{n} \sum_{i=0}^n i \frac{dc_i}{dt} \tag{5}$$

However, Equation (4) is not closed because we need to use an approximation scheme to write the transition rates γ_i in terms of the $c_i(t)$ variables. Note that the total fraction of susceptible



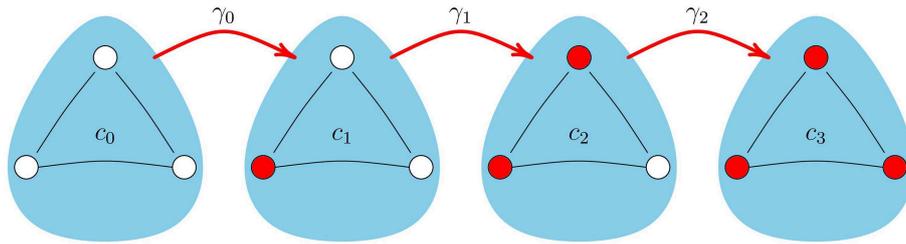


FIGURE 4 | Transition of a clique, with $n = 3$, between states. Note the transition rates γ_i are time-dependent quantities, given by Equation (9).

nodes in the network at time t is given by $1/n \sum_{i=0}^n (n-i)c_i$. We begin by calculating the probability, denoted Π_i , that at time t a chosen susceptible node is in a clique with i infected neighbors. This probability can be represented as the conditional probability $\Pr[i_{inf}|s] = \Pr[(i_{inf})\&(s)] / \Pr[s]$. The numerator is the joint probability of randomly selecting a susceptible node from a clique and the clique having i infected neighbors. To calculate this we first note that the probability of selecting a clique with i infected nodes is c_i and in a c_i clique the number of susceptible nodes is $(n-i)$. Thus, the probability of selecting a susceptible node from a c_i clique is $(n-i)/n$, yielding the required probability $\Pr[(i_{inf})\&(s)] = c_i(n-i)/n$. The denominator (the probability of selecting a susceptible node from a clique, $\Pr[s]$) can be obtained calculating the marginal distribution of s (i.e., by summing $c_i(n-i)/n$ over i) yielding $\sum_{i=0}^n c_i(n-i)/n$. Taking the ratio of the former to the latter yields the required probability

$$\Pi_i = \frac{(n-i)c_i}{\sum_{j=0}^n (n-j)c_j} \tag{6}$$

The probability distribution Π_i can be succinctly represented as a *probability generating function* (PGF) (see [23] for details), which is a polynomial function defined as

$$P(y) = \sum_{i=0}^n \Pi_i y^i; \tag{7}$$

note that the probabilities (Π_i) can be obtained in the usual way by repeated differentiation of the PGF:

$$\Pi_i = \frac{1}{i!} \left. \frac{d^i P}{dy^i} \right|_{y=0}. \tag{8}$$

This PGF provides a convenient method for calculating the probabilities inside a clique. However, a susceptible node in a chosen clique also receives exposures from infected nodes in other cliques, see **Figure 5**. Therefore, any approximation of γ_i needs to take into account not only the infected nodes inside a clique (the green area in **Figure 5**), but also the probability that the susceptible node comes into contact with infected nodes in its neighboring cliques (the blue area in **Figure 5**). Defining $\Pi_{i_e}^{m-1}$ as the probability that a susceptible node in a chosen clique has i_e infected neighbors in its other $m-1$ cliques, the probability

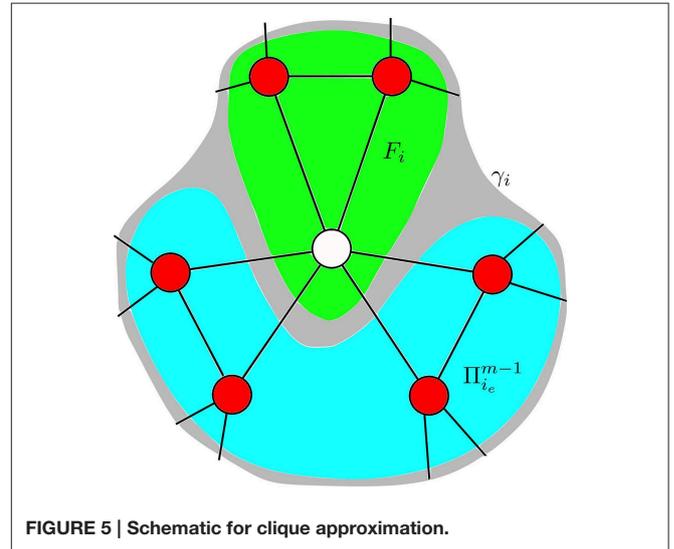


FIGURE 5 | Schematic for clique approximation.

distribution $\Pi_{i_e}^{m-1}$ has PGF $[P(y)]^{m-1}$. To approximate γ_i , we consider a clique with i infected nodes in it and look at one of the $n-i$ susceptible nodes to calculate the probability that this node transitions to the infected state. Such a transition changes the state of the clique, moving it from the c_i class to the c_{i+1} class. Consider the $m-1$ other cliques that the node is part of, letting i_e be the number of infected nodes present in the neighboring cliques, then the total number of infected neighbors is $i+i_e$ and the corresponding transition rate is F_{i+i_e} . (Here and henceforth, we write F_i in place of F_i^{CC}). Of course i_e can vary from 0 to $z-(n-1)$, therefore to approximate γ_i we weight F_{i+i_e} by the probability of observing i_e infected neighbors in the neighboring cliques, yielding:

$$\gamma_i = \sum_{i_e=0}^{z-n+1} \Pi_{i_e}^{m-1} F_{i+i_e}. \tag{9}$$

We assume that an initial fraction ρ_0 of randomly-chosen nodes are in the infected state at $t=0$. The probability that a clique contains i infected nodes at $t=0$ is therefore given by the binomial distribution:

$$c_i(t=0) = \binom{n}{i} (\rho_0)^i (1-\rho_0)^{n-i}, \tag{10}$$

which defines the initial conditions for the system given by Equation (4). With Equation (9) we can now solve Equation (4) numerically using the initial conditions (10) and thus calculate the total fraction of infected nodes at a given time for the networks of interest. In the next section we derive an early time approximation to the CA scheme to analytically find examples where quicker diffusion occurs on clustered networks than on the corresponding random network, similar to the results of Centola's experiments.

4.2. Linearization of the CA Model

In the previous section we derived the CA scheme, which captured the presence of clustering on clique-type networks. We wish to gain insight into the early spreading behavior produced by our complex contagion model (3) on the clustered networks outlined in Section 2. As previously mentioned, the CA scheme can be solved numerically using standard differential equation solvers, however it is also possible to find an approximate analytic solution to the early-time behavior. This is done by first perturbing the system (4) about a suitable fixed point and then linearizing the solution. The fixed point of interest is that corresponding to no infected nodes in the network ($c_0 = 1$ and $c_i = 0$ for $i \geq 1$). We perturb this fixed point by introducing a small positive parameter ϵ such that

$$\begin{aligned} c_0 &= 1 + \epsilon \tilde{c}_0, \\ c_i &= \epsilon \tilde{c}_i \quad \text{for } i > 0, \end{aligned} \tag{11}$$

where \tilde{c}_i are time-dependent quantities. Applying Equation (11) to the system of Equation (4) yields the perturbed system of equations

$$\begin{aligned} \epsilon \frac{d\tilde{c}_0}{dt} &= -n\gamma_0 - \epsilon n\gamma_0 \tilde{c}_0 \\ \epsilon \frac{d\tilde{c}_1}{dt} &= n\gamma_0 + \epsilon n\gamma_0 \tilde{c}_0 - \epsilon(n-1)\gamma_1 \tilde{c}_1 \\ \epsilon \frac{d\tilde{c}_i}{dt} &= \epsilon(n-i+1)\gamma_{i-1} \tilde{c}_{i-1} \\ &\quad - \epsilon(n-i)\gamma_i \tilde{c}_i \quad \text{for } i > 1. \end{aligned} \tag{12}$$

The γ_i 's require the approximation of $\Pi_{i_e}^{m-1}$, the probability that a susceptible node in a chosen clique has i_e infected neighbors in the remaining $m-1$ other cliques. These probabilities were built from the PGF defined by Equation (7) and applying the perturbation of Equation (11) to this results in

$$P(y) = 1 + \frac{\epsilon}{n} \sum_{i=0}^n \left((n-i)\tilde{c}_i y^i - (n-i)\tilde{c}_i \right) + \mathcal{O}(\epsilon^2), \tag{13}$$

where we are considering the asymptotic limit $\epsilon \rightarrow 0$ throughout this discussion and neglecting terms of order ϵ^2 and higher. We can find the PGF that corresponds to the distribution of probabilities for $\Pi_{i_e}^{m-1}$ by noting that

$$[P(y)]^{m-1} = 1 + \epsilon \frac{m-1}{n} \sum_{i=0}^n \left((n-i)\tilde{c}_i y^i - (n-i)\tilde{c}_i \right) + \mathcal{O}(\epsilon^2). \tag{14}$$

Next, we use Equation (14) to retrieve the required probabilities via the usual method of differentiation (Equation (8)). Using this relationship we find the first-order approximations

$$\begin{aligned} \Pi_0^{m-1} &\approx 1 - \epsilon \frac{(m-1)}{n} \sum_{i=0}^n (n-i)\tilde{c}_i, \\ \Pi_i^{m-1} &\approx \epsilon \frac{m-1}{n} (n-i)\tilde{c}_i \quad \text{for } i = 1 \text{ to } n. \end{aligned} \tag{15}$$

We are now able to approximate the γ_i 's by applying Equation (15) to Equation (9) and using the fact that $F_0 = 0$ (i.e., nodes require an infected neighbor before they can become infected), resulting in the following

$$\begin{aligned} \gamma_i &= \sum_{i_e=0}^n \Pi_{i_e}^{m-1} F_{i+i_e} \\ &= F_i - F_i \epsilon \frac{m-1}{n} \sum_{i_e=0}^n (n-i_e)\tilde{c}_{i_e} \\ &\quad + \epsilon \frac{m-1}{n} \sum_{i_e=1}^n (n-i_e)\tilde{c}_{i_e} F_{i+i_e} + \mathcal{O}(\epsilon^2). \end{aligned} \tag{16}$$

Inserting these rates into Equation (12) and noting that γ_0 is $\mathcal{O}(\epsilon)$ while $\gamma_i = F_i + \mathcal{O}(\epsilon)$ for $i \geq 1$, we obtain the linearization of the CA system:

$$\begin{aligned} \frac{d\tilde{c}_0}{dt} &= -(m-1) \sum_{i_e=1}^n (n-i_e)\tilde{c}_{i_e} F_{i_e} + \mathcal{O}(\epsilon), \\ \frac{d\tilde{c}_1}{dt} &= (m-1) \sum_{i_e=1}^n (n-i_e)\tilde{c}_{i_e} F_{i_e} - (n-1)F_1 \tilde{c}_1 + \mathcal{O}(\epsilon), \\ \frac{d\tilde{c}_i}{dt} &= (n-i+1)F_{i-1} \tilde{c}_{i-1} - (n-i)F_i \tilde{c}_i + \mathcal{O}(\epsilon) \quad \text{for } i > 1. \end{aligned} \tag{17}$$

Now we have a system of equations that describes the early spreading behavior for a general transition rate function F_i . We want to use this to find a linearized solution $(\rho_i(t))$ that approximates the behavior of the CA scheme. Let $\mathbf{C} = (\tilde{c}_0, \dots, \tilde{c}_{n-1})^T$ and further define $d\mathbf{C}/dt = \mathbf{f}(\mathbf{C}, t)$. The linearized system (17) is defined by

$$\frac{d\mathbf{C}}{dt} = \mathbf{J}\mathbf{C}, \tag{18}$$

where \mathbf{J} is the $n \times n$ Jacobian matrix with element $\partial \mathbf{f}_i / \partial \mathbf{C}_j$ in the i th row and the j th column. Note that the \tilde{c}_n variable does not feature in our calculation of \mathbf{J} because it is fully determined by the relationship $\tilde{c}_n = 1 - \sum_{i=0}^{n-1} \tilde{c}_i$. The general solution of systems like (18) typically can be written as

$$\mathbf{C}(t) = \sum_{j=0}^{n-1} \xi_j e^{\lambda_j t} \mathbf{u}_j, \tag{19}$$

where ξ_j is a constant, λ_j is the eigenvalue and \mathbf{u}_j is the corresponding eigenvector of \mathbf{J} [24]. The constants ξ_j can be

calculated by using the initial conditions $\mathbf{C}(t = 0) = \sum_{j=0}^n \xi_j \mathbf{u}_j$ (refer to Equation (10) for initial conditions for the system). The fixed point that we considered was $c_0 = 1$ and $c_i = 0$ for $i > 0$, where there were no infected nodes on the network. Our linearized solution is therefore valid for small perturbations from this, i.e., when the initial fraction of infected nodes is small ($\rho(t)$ is small and $\mathcal{O}(\rho_0^2)$ terms are negligible). The linearized approximation to the total fraction of infected nodes at time t is given by

$$\rho(t) \approx \rho_l(t) = \frac{1}{n} \sum_{i=0}^n i \tilde{c}_i(t). \quad (20)$$

This formulation now allows us to examine the early-time spreading behavior that is produced by our complex contagion model (and see Appendix B for a simple worked example). It is also possible to find the level of social reinforcement β for which a clustered network will propagate a complex contagion faster than a random network. The largest eigenvalue of the Jacobian matrix (which we denote λ_{max}) appearing in the linearization Equation (19) provides the largest contribution to the early-time growth of Equation (20), and so to $\rho(t)$. Thus, by comparing the λ_{max} value for each network for a given β and noting which network has the larger value, we can infer the case where the complex contagion will diffuse faster, at least at early times. This will be used in the following section in conjunction with the full CA scheme and the linearized solution to examine the complex contagion model on networks with various levels of clustering.

5. Results

In Section 4.1 we described the *clique approximation* (CA) scheme that we use to account for the presence of clustering in clique-type networks for monotone binary-state dynamics. We also linearized the CA scheme to approximate the early-time spreading behavior (Section 4.2). In this section, we compare the accuracy of the full CA scheme and the linearized approximation to Monte Carlo (MC) simulations of the complex contagion model given by Equation (3) (for details on simulations please refer to Appendix C). This allows us to establish the accuracy of both the CA scheme and its linearized approximation across clique-type networks and varying level of social reinforcement (as parameterized by β).

Recall that we consider three z -regular network topologies with degree 6 (refer back to **Figure 1**). First, a random network ($n = 2$ and $m = 6$), which has the lowest density of triangles ($C_\Delta = 0$), then a moderately clustered network where each clique has three nodes and each node is part of three cliques ($C_\Delta = 0.2$), and lastly, a highly clustered network where each clique has four nodes and each node is a part of 2 cliques ($C_\Delta = 0.4$). **Figure 6** presents the results across the three topologies that we consider and for two values of β . The CA method clearly provides a highly accurate approximation to $\rho(t)$ across the three network topologies. The linearized approximation of the CA scheme also provides accurate approximations for the early-time growth of $\rho(t)$. However, once the fraction of infected nodes becomes large

during the later stages of spreading the approximation begins to break down.

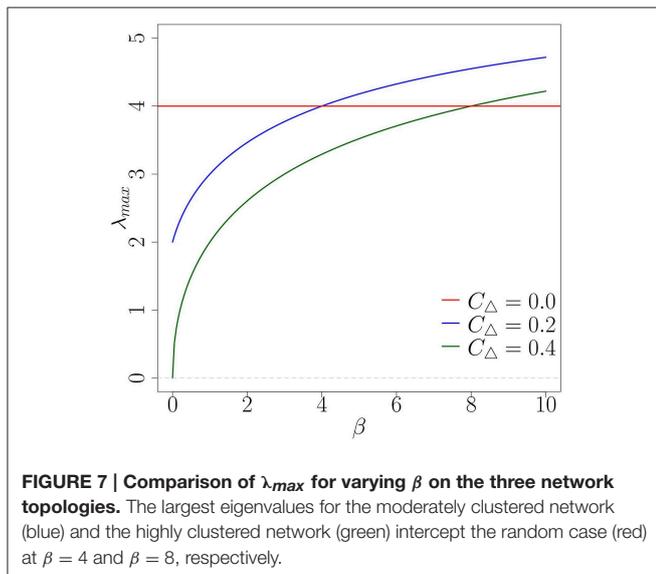
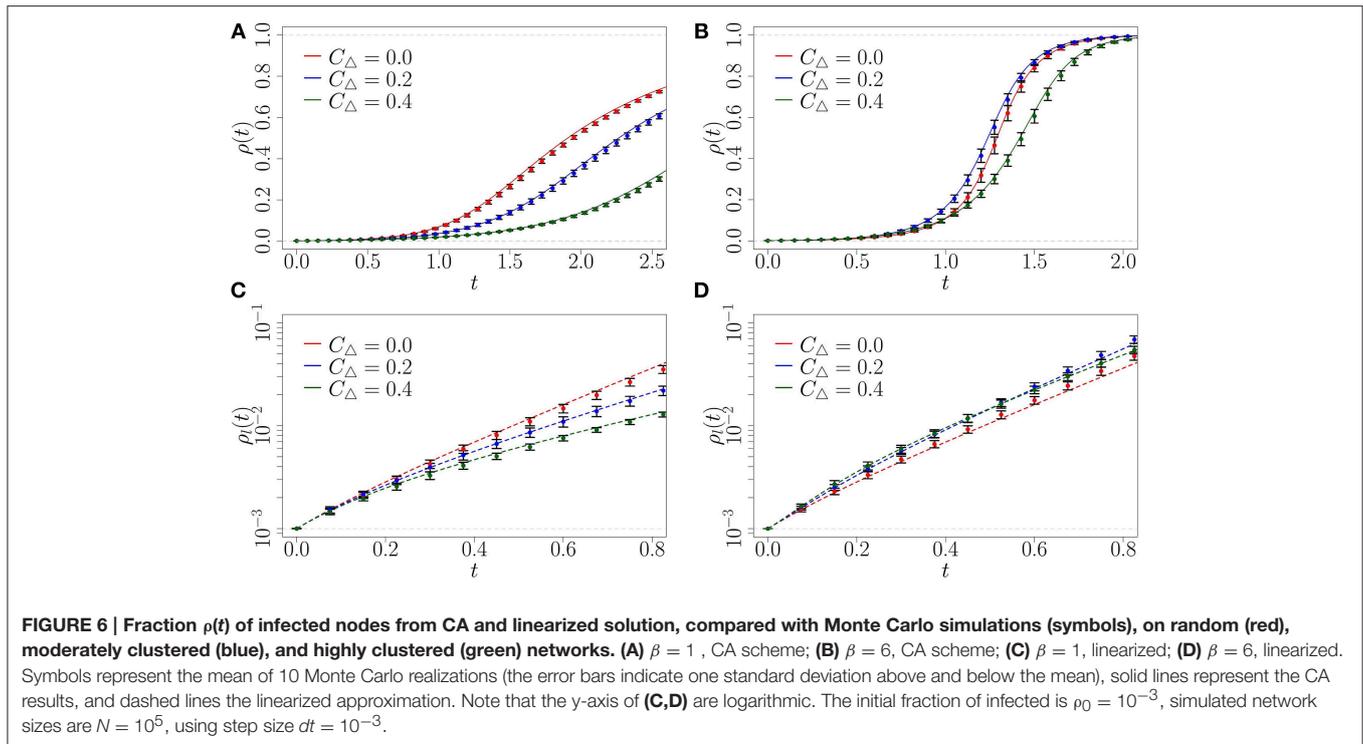
Now we examine the spreading behavior that our complex contagion model F_i^{CC} produces on clustered networks. In the definition of F_i^{CC} the parameter β is the rate at which a susceptible node will become infected if more than one of its neighbors is infected. As β increases we expect the infection to spread faster on the two clustered networks than on the random network (at least at early times) because of the existence of reinforcement signals from triangles.

For comparison, we consider $\beta = 1$, meaning that a susceptible node with one infected neighbor has the same infection rate as a susceptible node with multiple infected neighbors. From **Figures 6A,C** we see that in this case the behavior spreads fastest on the random ($C_\Delta = 0$) network, because the random network allows the maximum number of unique exposures from newly infected nodes.

However, for larger values of β it becomes more advantageous for early-stage spreading to have a non-zero density of triangles than a tree-like structure in the local topology. By increasing β to 6 we find this is the case (see **Figure 6D**). Note that at early times (before $t = 1$) the random network consistently infects a lower fraction of the population than the clustered networks; we analyze this phenomenon further below.

Empirical observations of spreading behavior on networks shows that typically only a small fraction of the total network ever adopts a behavior. Centola [15] observed that the average percentage of the network that adopted was 38 and 53% for the random and clustered networks respectively. The networks used in his experiments were relatively small, with $N \leq 144$ nodes. If our complex contagion model is reflective of the spreading behavior in real life contagions we should observe the same behavior for small $\rho(t)$ which corresponds to the early-time behavior (which we consider in **Figure 8**). Before we examine this in detail we calculate the critical reinforcement levels for which we expect clustered networks to produce faster early time spreading than the random network (at least in the limit of very large network size, $N \rightarrow \infty$, for which our approximations are valid).

As mentioned in Section 4.2, by finding the network topology with the largest λ_{max} for a given β we can identify which network will produce the fastest diffusion of an early-stage complex contagion. For the random network topology ($C_\Delta = 0$) the largest eigenvalue is $\lambda_{max} = 4$. For the moderately clustered network topologies the largest eigenvalue is $\lambda_{max} = 1/2(2 - \beta + \sqrt{4 + 20\beta + \beta^2})$, while the highly clustered network topology has largest eigenvalue $\lambda_{max} = 1/2(-\beta + \sqrt{\beta}\sqrt{24 + \beta})$. By plotting how these vary with β we can identify the level of social reinforcement required to produce faster spreading on the clustered networks than on the random network at early time. From **Figure 7** we note that for $\beta > 4$ (respectively, $\beta > 8$) the moderately (highly) clustered network should produce faster diffusion than the random network. The main limitation of the predicted critical β 's is that the exponential growth rate λ_{max} must dominate in Equation (19) over a sufficient range for its contributions to become pronounced. To obtain this behavior ρ_0 must be very small. This ensures that the initial transient behavior



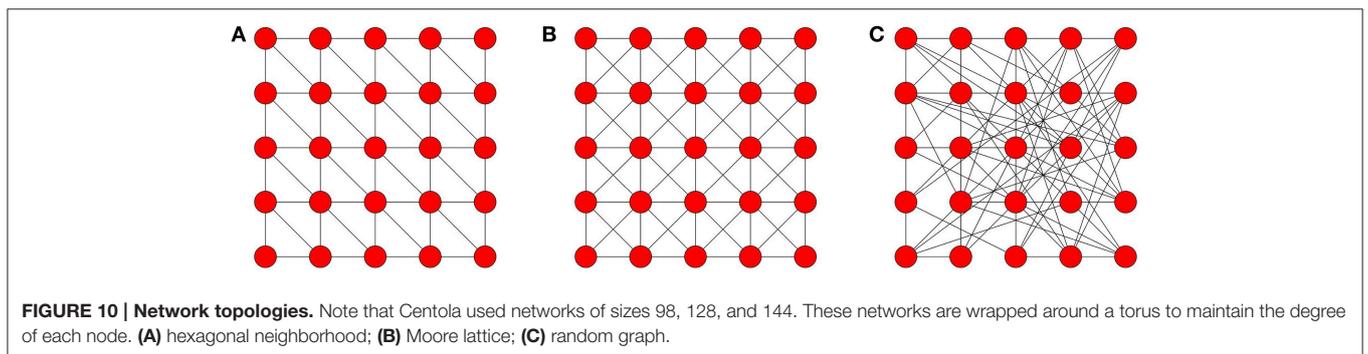
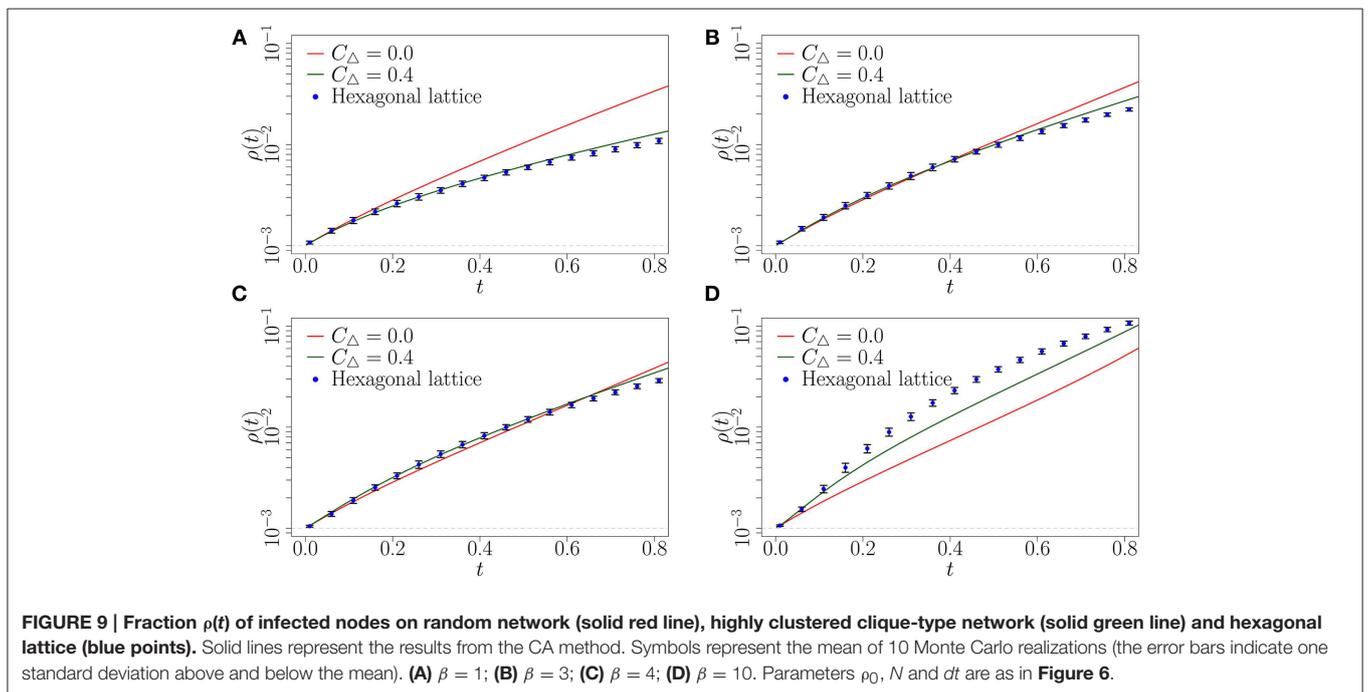
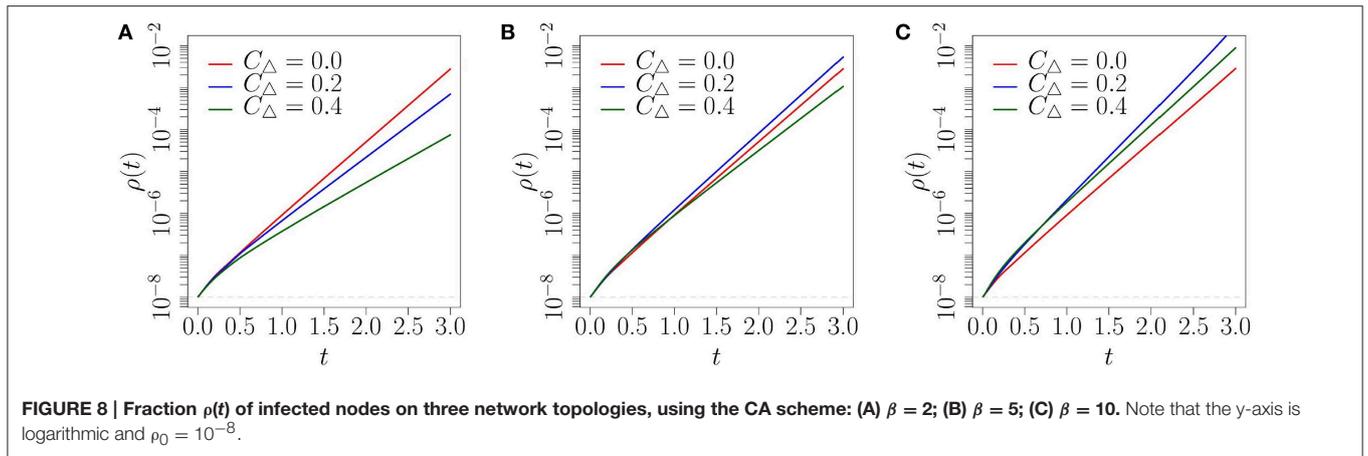
(the contributions from the other eigenvalues) dies off quickly and the exponential growth at rate λ_{max} dominates. Therefore, in **Figure 8** we show the predicted fraction of infected nodes from the full CA model at early stages for a very small fraction of infected nodes, $\rho_0 = 10^{-8}$, which would correspond to a very large network.

Similar to what we observed in **Figure 6**, the level of social reinforcement dictates how fast the diffusion spreads on each network at early times in **Figure 8**. We also observe that the

order of the networks that provide the fastest diffusion is well reflected by the comparison of each network's λ_{max} illustrated in **Figure 7**. More specifically, in **Figure 8A** where $\beta = 2$, we see that the level of social reinforcement is not high enough to cause faster spreading on the clustered networks than on the corresponding random network. Increasing β to 5 we observe faster spreading on the moderately clustered network than on the random network, with the highly clustered network producing the slowest diffusion (see **Figure 8B**). Increasing β further to 10 we observe faster spreading of both clustered networks over the random network (**Figure 8C**), again in accordance with what is expected from **Figure 7**. Although the critical levels of social reinforcement predicted in **Figure 7** are accurate for $\rho_0 \ll 1$, qualitatively similar behavior is produced for larger values of ρ_0 (refer to **Figure 6**), but with stronger influence of initial transients.

Finally, we show for completeness that our complex contagion model can produce faster spreading on a hexagonal lattice compared with a random network, which mimics Centola's experimental setup (see Appendix A for details). The topology of the hexagonal lattice is illustrated in **Figure 10A**, and it has clustering coefficient of 0.4. We simulate the complex contagion on this network using the Monte Carlo method on large networks ($N = 10^5$) with the hexagonal lattice structure.

The results of the simulations are compared to the expected diffusion on a random and highly clustered network of the same degree ($z = 6$) using the CA method (see **Figure 9**). We find similar results to those noted in the analysis of the clique-type networks. For low levels of social reinforcement ($\beta \leq 3$) the random network provides the fastest spreading



(see Figures 9A,B). However, when β is increased to 4 we observe faster diffusion on the hexagonal lattice and the highly clustered network than on the corresponding random network

(see Figure 9C). Both the hexagonal lattice and the highly clustered network appear to have roughly the same critical level of social reinforcement for this initial fraction of infected nodes

($\rho_0 = 10^{-3}$). By increasing β further to 10 (see **Figure 9D**) we find that the hexagonal lattice provides the fastest diffusion at early time. This is interesting as both the highly clustered network and the hexagonal lattice have the same clustering coefficient of $C_\Delta = 0.4$. This is explained by the difference in structure between the hexagonal lattice and the highly clustered network. The hexagonal lattice has a higher density of cycles of length greater than 3. The highly clustered network on the other hand has a lower density of cycles of length greater than 3 as each node is randomly connected to each clique. This results in faster spreading on the hexagonal lattice at early time than on the highly clustered network due to the increased chance of a susceptible node receiving multiple exposures from infected nodes. This qualitatively reproduces the pattern of spreading behavior observed by Centola, where for a sufficient level of social reinforcement it is possible to produce faster spreading on clustered networks than a random network of the same degree. In the next section we conclude the paper with a summary, some comments on the results and provide possible directions for future research.

6. Conclusion

In this paper we aimed to model—in an analytically tractable fashion—the spreading of behaviors such as the adoption of new innovations. Such spreading processes are influenced by the social networks that connect people. Centola performed an experiment where he tracked the diffusion of such behavior (the use of a health forum) across artificially created networks [15]. These networks allowed him to control the level of clustering (density of cycles of length three) in the local topology and to isolate its effect on how the behavior diffused (refer to **Figure 10**). He observed that nodes that received multiple reinforcing signals had a higher propensity to adopt compared to those that only received one signal, which was much more beneficial to the spreading of the contagion on the clustered networks. This resulted in the contagion spreading farther and faster on the clustered-lattices than on the corresponding random networks.

Our goal was to find a suitably simple characterization for complex contagion that remained amenable to analysis. We proposed modeling the complex contagion using monotone binary-state dynamics with the transition rate function defined by F_i^{CC} (see Section 3). Each node is either susceptible (has not yet adopted) or infected (adopted). This simple characterization proved to be quite effective in enabling us to obtain analytical insight. We compared the spreading behavior produced by the complex contagion model across three topologies with varying levels of clustering (see **Figure 1** for the topologies and **Figure 6** for results). By varying the propensity for a node to become infected given multiple infected neighbors we were able to produce faster spreading on clustered networks than on the random network, which is qualitatively similar

behavior to that observed by Centola (**Figure 6B**). We also showed, via simulation, that our complex contagion model could produce similar spreading behavior between a hexagonal lattice and comparable random network as the previously mentioned analytic results for the clique-type networks (see **Figure 9**).

None of these results could have been obtained without tackling the problem of approximating monotone binary-state dynamics on clustered networks. As described in Section 4.1, standard approximation schemes (mean-field and pair approximation) perform poorly in the presence of clustering. They are heavily dependent on the assumption that the network is locally tree-like (that is no cycles of length three in the network). However, the use of clustered networks is crucial to the examination of the complex contagion model, as the presence of triangles are central to the social reinforcement mechanism that we wished to examine. This necessitated the development of the CA method which accurately accounted for the effects of clustering in the local topology of the clique-based networks we examined (see Section 4 for details). The CA method proved to be highly accurate for these types of topologies. A linearized approximation to the early-time spreading behavior of the complex contagion model was obtained. Using this we were able to calculate critical levels of social reinforcement required for the contagion to spread faster on clustered networks than on the corresponding random network (refer to **Figure 7**).

The characterization of a complex contagion spreading process by a single-parameter function in Equation (3) provided a suitable balance between simplicity and realistic behavior. However, the approximation scheme we develop is applicable to any F_i function, and so more realistic models can easily be examined in this framework. Further examination of more realistic characterization of complex contagions should also be developed, for example including a time-decay in the memory of each node. It is reasonable to assume that the true mechanism that governs complex contagion depends on the interplay between the strength of social reinforcement and also temporal effects, such as the timing between exposures.

Author Contributions

JG and PF designed the research and developed the approximation schemes; DOS and GOK performed the calculations and numerical simulations; DOS led the writing of the paper.

Acknowledgments

This work was partly funded by Science Foundation Ireland (awards 11/PI/1026 and 12/IA/I683), the Irish Research Council (award GOIPG/2014/887) and by the European Commission through FET-Proactive project PLEXMATH (FP7-ICT-2011-8; grant number 317614).

References

- Newman ME. Properties of highly clustered networks. *Phys Rev E* (2003) **68**:026121. doi: 10.1103/PhysRevE.68.026121
- Ugander J, Backstrom L, Marlow C, Kleinberg J. Structural diversity in social contagion. *Proc Natl Acad Sci USA*. (2012) **109**:5962–6. doi: 10.1073/pnas.1116502109
- Broder A, Kumar R, Maghoul F, Raghavan P, Rajagopalan S, Stata R, et al. Graph structure in the web. *Comput Netw.* (2000) **33**:309–20. doi: 10.1016/S1389-1286(00)00083-9
- Maslov S, Sneppen K. Specificity and stability in topology of protein networks. *Science* (2002) **296**: 910–3. doi: 10.1126/science.1065103
- Aral S, Muchnik L, Sundararajan A. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proc Natl Acad Sci USA*. (2009) **106**:21544–9. doi: 10.1073/pnas.0908800106
- Baños RA, Borge-Holthoefer J, Moreno Y. The role of hidden influentials in the diffusion of online information cascades. *EPJ Data Sci.* (2013) **2**:1–16. doi: 10.1140/epjds18
- Nematzadeh A, Ferrara E, Flammini A, Ahn YY. Optimal network modularity for information diffusion. *Phys Rev Lett.* (2014) **113**:088701. doi: 10.1103/PhysRevLett.113.088701
- Borge-Holthoefer J, Baños RA, González-Bailón S, Moreno Y. Cascading behaviour in complex socio-technical networks. *J Complex Netw.* (2013) **1**:3–24. doi: 10.1093/comnet/cnt006
- Centola D, Macy M. Complex contagions and the weakness of long ties. *Am J Sociol.* (2007) **113**:702–34. doi: 10.1086/521848
- Kitsak M, Gallos LK, Havlin S, Liljeros F, Muchnik L, Stanley HE, et al. Identification of influential spreaders in complex networks. *Nat Phys.* (2010) **6**:888–93. doi: 10.1038/nphys1746
- Porter MA, Gleeson JP. Dynamical systems on networks: a tutorial (2014). arXiv preprint arXiv:1403.7663.
- Newman ME, Park J. Why social networks are different from other types of networks. *Phys Rev E* (2003) **68**:036122. doi: 10.1103/PhysRevE.68.036122
- Watts DJ, Strogatz SH. Collective dynamics of small-world networks. *Nature* (1998) **393**:440–2. doi: 10.1038/30918
- Shirley MD, Rushton SP. The impacts of network topology on disease spread. *Ecol Complex.* (2005) **2**:287–99. doi: 10.1016/j.ecocom.2005.04.005
- Centola D. The spread of behavior in an online social network experiment. *Science* (2010) **329**:1194–7. doi: 10.1126/science.1185231
- Lü L, Chen DB, Zhou T. The small world yields the most effective information spreading. *N J Phys.* (2011) **13**:123005. doi: 10.1088/1367-2630/13/12/123005
- Pastor-Satorras R, Vespignani A. Epidemic spreading in scale-free networks. *Phys Rev Lett.* (2001) **86**:3200. doi: 10.1103/PhysRevLett.86.3200
- Gleeson JP. Binary-state dynamics on complex networks: pair approximation and beyond. *Phys Rev X* (2013) **3**:021004. doi: 10.1103/PhysRevX.3.021004
- Hébert-Dufresne L, Noël PA, Marceau V, Allard A, Dubé LJ. Propagation dynamics on networks featuring complex topologies. *Phys Rev E* (2010) **82**:036115. doi: 10.1103/PhysRevE.82.036115
- Newman MEJ. *Networks: An Introduction*. New York, NY: Oxford University Press (2010). doi: 10.1093/acprof:oso/9780199206650.001.0001
- Newman MEJ. The structure and function of complex networks. *SIAM Rev.* (2003) **45**:167–256. doi: 10.1137/S003614450342480
- Miller JC. Percolation and epidemics in random clustered networks. *Phys Rev E* (2009) **80**:020901. doi: 10.1103/PhysRevE.80.020901
- Johnson NL, Kemp AW, Kotz S. *Univariate Discrete Distributions*. Vol. 444. Hoboken, NJ: John Wiley & Sons (2005). doi: 10.1002/0471715816
- Drazin PG. *Nonlinear Systems*. Vol. 10. Cambridge, UK: Cambridge University Press (1992).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 O'Sullivan, O'Keefe, Fennell and Gleeson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Appendix

A. Centola's Experimental Design

Centola investigated the effects of network structure on the spread of behavior through artificially structured online communities. These networks were carefully created to allow for direct comparison between random graphs (low C_Δ) and clustered lattices (high C_Δ).

In his experiment each network represents the connections of an artificially created on-line health community. Each participant created an anonymous on-line profile, they were then linked to other participants according to a predefined topology (see **Figure 10** for examples). Each participant could not communicate with each other directly but were informed of their activities. Participants made decisions on whether or not to adopt a behavior based on their neighbors' activity, in this case the registration to a health forum. The diffusion process was initiated by selecting a random seed node, which signaled (via an automatically sent email) its neighbors, encouraging them to register for the forum. Every time a participant adopted the behavior (registered to the forum), messages were sent to his network neighbors. As the number of nearest neighbors that registered increased, the participant received more signals. Several trials were conducted on two different clustered topologies and two unclustered topologies.

The Hexagonal lattice and Moore lattice corresponded to the clustered topologies used, each having a fixed degree of six and eight for all nodes with a clustering coefficient of 0.4 and 0.43 (see **Figures 10A,B**), respectively. Each clustered topology was then compared to a random network, where each node has the same degree but the links were randomly assigned, as illustrated in **Figure 10C** (a random network with fixed degree). Random graphs have the lowest clustering coefficients of all the graphs that Centola used.

B. Linearization - Worked Examples

In this section we provide a simple example of the linearized approximation of the CA scheme described in Section 4.2. The spreading dynamics that we will approximate is the behavior of our complex contagion model defined in Section 3 on a z -regular random network (zero clustering case). Recall that random networks can be described by $n = 2$ and $m = 6$, where each node in the network has degree six (see **Figure 1**). Applying these values to the linearized system of Equation (17) we have

$$\begin{aligned} \frac{d\tilde{c}_0}{dt} &= -5 \sum_{i_e=1}^2 (2 - i_e) \tilde{c}_{i_e} F_{i_e}^{CC} + \mathcal{O}(\epsilon), \\ \frac{d\tilde{c}_1}{dt} &= 5 \sum_{i_e=1}^2 (2 - i_e) \tilde{c}_{i_e} F_{i_e}^{CC} - (2 - 1) F_1^{CC} \tilde{c}_1 + \mathcal{O}(\epsilon). \end{aligned} \tag{A1}$$

Recall that we do not require the \tilde{c}_n variable as it is fully determined by the other \tilde{c}_i variables. The next stage in this

approximation is to compute the Jacobian matrix for this system:

$$\mathbf{J} = \begin{pmatrix} 0 & -5 \\ 0 & 4 \end{pmatrix}. \tag{A2}$$

Matrix (A2) has eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 0$, each with associated eigenvectors $\mathbf{u}_1 = (-5, 4)^T$ and $\mathbf{u}_2 = (1, 0)^T$, respectively. Applying these to Equation (19), we obtain

$$\mathbf{C}(t) = \xi_1 e^{\lambda_1 t} \mathbf{u}_1 + \xi_2 e^{\lambda_2 t} \mathbf{u}_2. \tag{A3}$$

The constants (ξ_1 and ξ_2) can be easily obtained by noting that $\mathbf{C}(t = 0) = \sum_{j=0}^n \xi_j \mathbf{u}_j$ must equal the initial conditions of Equation (10). We ignore $\mathcal{O}(\rho_0^2)$ terms, which yields $\xi_1 = \rho_0/2$ and $\xi_2 = 1 - \rho_0/2$. With these constants we are able to calculate the linearized approximation to the fraction of infected nodes on a z -regular random network given by Equation (20) as

$$\rho_i(t) = \frac{3}{2} \rho_0 e^{4t} - \frac{\rho_0}{2}. \tag{A4}$$

Notably, in this simple example we find that $\beta (F_i^{CC}$ for $i \geq 2$) does not feature in the result, this is a direct consequence of the assumption that the local topology generated by $n = 2$ and $m = 6$ is locally tree-like. Equation (A4) provides an accurate approximation to the early-time spreading behavior of a complex contagion on a tree-like network of degree 6, provided that the initial fraction of infected nodes is small (see **Figure 6**).

C. Simulation Method

To simulate monotone binary-state dynamics we use Monte Carlo (MC) simulation. To represent the network in the simulations we use an *adjacency matrix* \mathbf{A} , where

$$A_{ij} = \begin{cases} 1 & \text{if there is a link between nodes } i \text{ and } j, \\ 0 & \text{otherwise,} \end{cases} \tag{A5}$$

defines an $N \times N$ matrix, where N is the number of nodes. Given this matrix we know the connections between nodes. We track the state of each node using the vector \mathbf{v} (a $N \times 1$ vector). The element v_i is 0 if the node is susceptible and 1 if infected. To initialize the simulation we randomly assign a fraction ρ_0 of nodes to the infected state at time 0. We wish to simulate the dynamics for the complex contagion model defined in Section 3 by Equation (3). The transition rates in these models depend on the number of infected neighbors of a node. Let $\boldsymbol{\eta}$ be an $N \times 1$ dimensional vector where η_i is the number of infected neighbors of node i . The vector of η_i values can be easily calculated using the matrix multiplication

$$\boldsymbol{\eta} = \mathbf{A}\mathbf{v}. \tag{A6}$$

The probability that node i will change state is given by $p = F_{\eta_i} dt$, where F_{η_i} is the transition rate for node i (which has η_i infected neighbors). This gives the update rule for the state of

node i , where $v_i = 1$ if $p > u$, for u drawn from a uniform distribution on $[0, 1]$. The fraction of infected nodes is then updated ($\rho(t + dt) = \frac{1}{N} \sum \mathbf{v}$) and time, t , is advanced by dt . These steps are repeated until either $\rho(t) = 1$ or until a maximum time is reached (t_{max}). This process yields one realization of the

dynamics, it is repeated M times (the number of MC realizations) and the ensemble-average fraction of infected nodes is calculated to approximate the expected behavior of the dynamics. The parameters used for simulations are as follows unless otherwise stated: $N = 10^5$, $\rho_0 = 10^{-3}$, $t_{max} = 3$, $dt = 10^{-3}$ and $M = 10$.

Information transfer in community structured multiplex networks

Albert Solé-Ribalta*, Clara Granell, Sergio Gómez and Alex Arenas

Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Tarragona, Spain

OPEN ACCESS

Edited by:

Taha Yasseri,
University of Oxford, UK

Reviewed by:

Renaud Lambiotte,
University of Namur, Belgium
Nuno A. M. Araújo,
Universidade de Lisboa, Portugal

*Correspondence:

Albert Solé-Ribalta,
Departament d'Enginyeria Informàtica
i Matemàtiques, Universitat Rovira i
Virgili, Avinguda Països Catalans 26,
43007 Tarragona, Spain
albert.sole@urv.cat

Specialty section:

This article was submitted to
Interdisciplinary Physics,
a section of the journal
Frontiers in Physics

Received: 30 June 2015

Accepted: 28 July 2015

Published: 18 August 2015

Citation:

Solé-Ribalta A, Granell C, Gómez S
and Arenas A (2015) Information
transfer in community structured
multiplex networks. *Front. Phys.* 3:61.
doi: 10.3389/fphy.2015.00061

The study of complex networks that account for different types of interactions has become a subject of interest in the last few years, specially because its representational power in the description of users interactions in diverse online social platforms (Facebook, Twitter, Instagram, etc.). The mathematical description of these interacting networks has been coined under the name of multilayer networks, where each layer accounts for a type of interaction. It has been shown that diffusive processes on top of these networks present a phenomenology that cannot be explained by the naive superposition of single layer diffusive phenomena but require the whole structure of interconnected layers. Nevertheless, the description of diffusive phenomena on multilayer networks has obviated the fact that social networks have strong mesoscopic structure represented by different communities of individuals driven by common interests, or any other social aspect. In this work, we study the transfer of information in multilayer networks with community structure. The final goal is to understand and quantify, if the existence of well-defined community structure at the level of individual layers, together with the multilayer structure of the whole network, enhances or deteriorates the diffusion of packets of information.

Keywords: complex networks, information diffusion, multivariate analysis, community structure, centrality

1. Introduction

The study of transport properties of networks is becoming increasingly important due to the constantly growing amount of information and commodities being transferred through them. A particular focus of these studies is how to make the capacity of the diffusion of information in the network maximal while minimizing the delivery time. In the basic approach information is formed by units, the “packets,” and the handling of information for processing and distribution takes finite time. Both network packet routing strategies and network topology play an essential role in networks’ traffic flow. In realistic settings, like online social networks, the knowledge that any one has about the topology of the network is limited to its local area of influence. Consequently, much of the focus in recent studies has been on “searchability,” the process of sending information to a target when the trajectory to reach the target is unknown. Moreover, given the limited capability of nodes to handle information packets and redistribute them, the problem of congestion arises [1–3]. It has been observed, both in real world networks and in model communication networks, that the network flow collapses when the load (number of packets to be processed) is above a certain threshold [3].

In general, most real and engineered systems include multiple subsystems and layers of connectivity, and it is important to take such features into account when trying to obtain a complete understanding of them. It is thus necessary to generalize the “traditional” network theory

to multilayer systems in a comprehensive fashion [4, 5]. Generally speaking, up to now, the description of networks so far has been developed using a single and combined snapshot of the connectivity, which is a reflection of instantaneous interactions or accumulated interactions in a certain time window. This description is limited when trying to understand the intricate variability of real complex systems, which contain many different time scales [6] and coexisting structural patterns forming the real network of interactions [7]. This is the case of e-social networks that are constantly changing [8], having some connections with very short lifetime and others that are persistent. Interest groups [9] are constantly being developed and growing, and individual nodes participate through different interests at the same time. An accurate description of such complexity should take into account these differences of interactions and their evolution through time. In the last couple of years, the scientific community on networks theory has focused on this issue and proposed a solution that has been commonly referred to as the multiplex network structure [4].

General flows on multiplex networks have also been in the focus of network scientists [10–17], and the consequences of having such topologies have been shown to be far from trivial. For example, in Solé-Ribalta et al. [18] the authors found that a general diffusive process on top of the multiplex structure is able to speed up the less diffusive of the layers. It could also give rise to a super-diffusion process thus enhancing the diffusion of both layers. This striking result appears when the diffusion between the layers of the multiplex is faster than that occurring within each of the layers. These consequences are also observed in the discrete representation of diffusion by random walkers [17], and have explicit consequences on the navigability of the multiplex structure.

Here we fix our attention in the process of information transfer on top of multiplex networks. Specifically, we aim at determining the structural effects of a multiplex network endowing community structure, i.e., modular at each layer, on the dynamics of information transfer. To this end we have investigated a particular set up in which multiplex networks are built connecting different modular networks, and determining analytically the onset of congestion in the information flow. Our results reveal that when the community structure of the different layers is equivalent and the communities overlap, the multiplex offers higher resilience to congestion and consequently the system may improve information transfer compared to the individual layers. On the other side when the community structure is considerably different and communities still overlap the multiplex structure offers a balancing environment where the efficiency of the system is averaged. On the intermediate situation, that is community structure is similar in both layers and communities overlap, the effect of the multiplex structure is devastating and hinders information transfer by reducing the onset of congestion in the system.

2. Materials and Methods

The proposed dynamical model considers that information flows through networks in atomic and discrete packets that are sent

from an origin node to a destination node. Each node is an independent agent that can store as many packets as necessary. However, to have a realistic picture of communication we must assume that the nodes have a finite capacity to process and deliver packets. That is, a node will take longer to deliver two packets than just one. This physical constraint of the agents on delivering information can derive in network congestion. When the amount of information a particular agent receives is too large, it is not able to handle all the packets and some of them remain undelivered for extremely long periods of time. In this study, the interest is focused on when congestion occurs depending on the topology of the multiplex network, in particular, in relation to its community structure.

2.1. Dynamics of Information Transfer

The dynamics of the model is as follows. At each time step t , information packets are created at every node with rate ρ (injection rate). Therefore, ρ is the control parameter: small values of ρ correspond to low density of packets and high values of ρ correspond to the generation of a large amount of packets. When a new packet is created, a destination node, different from the origin, is chosen (uniformly) at randomly in the network. Thus, during the following time steps $t + 1, t + 2, \dots, t + T$, the packet travels toward its destination. Once the packet reaches the destination node, it is delivered and disappears from the network.

The time that a packet remains in the network is related not only to the length of the path between the source and the target nodes, but also to the volume of packets that share its path. Nodes with high loads, i.e., high volume of accumulated packets, will take longer to deliver packets or, in other words, it will take more time steps for packets to cross regions of the network that are highly congested. We assume, without loss of generality, that nodes can handle only one packet per time step (i.e., the delivery rate is $\tau = 1$), and undelivered packets are stored in a first-in-first-out queue attached to each node. The paths followed by packets between source and destination nodes are decided using a routing strategy, being shortest paths and random walks the most prominent strategies. It is important to note, however, that the model is not deterministic. For example, there may be several shortest paths between two nodes, one of them chosen randomly in the delivery of the corresponding packet. Moreover, the order in which packets are stored in the queues when several of them arrive in the same time step is undefined.

Previous work on single layer networks [3] shows that for low values of the injection rate of packets ρ there is no accumulation of packets at any node in the network. Moreover, it is stated that the number of packets that arrive at node i is, on average, $\rho B_i / (N - 1)$, where B_i is the effective betweenness of node i and N the number of nodes in the network. The effective betweenness is defined as the ratio between the number of paths that pass through node i , and the total number of paths traversing the network between any pair of nodes [19].

The onset of congestion is reached when a node receives more packets than it can deliver per time step, i.e., $\rho B_i / (N - 1) > 1$. Therefore, the first node that collapses (i^*) is the one with largest effective betweenness ($B_{i^*} = \max_i(B_i)$), and the maximum injection rate for which the network is congestion free, the critical

injection rate ρ_c is given by

$$\rho_c = \frac{N-1}{B_{i^*}}. \quad (1)$$

The rest of the nodes will collapse with larger injection rates. However, up to now, it is not known how to analytically compute their critical injection rates since they not only depend on the topological betweenness but also on the overall network congestion.

In the generalization of the routing dynamics exposed to multiplex networks, the average number of packets arriving to node i in layer α is $\rho L B_{i\alpha} / (N-1)$, where L is the number of layers of the multiplex network and $B_{i\alpha}$ is the effective betweenness of node i in layer α . Thus, the critical injection rate also depends on the effective betweenness, which encapsulates the routing strategy and the topology of the network:

$$\rho_c = \frac{(N-1)/L}{\max_{i,\alpha}(B_{i\alpha})}. \quad (2)$$

Next, we extend the concept of betweenness to multiplex networks allowing the computation of the onset of congestion.

2.2. Computation of betweenness in the Multiplex

The extension of any centrality measures to multiplex networks is not straightforward and requires special care. In many situations several extensions are possible and the choice of it may depend on the problem at hand. Many attempts have been done to extend single layer centrality measures to the multiplex framework [20–22]. Here, we follow the line described in De Domenico et al. [23], which is mathematically grounded on the tensorial formalism for multilayer networks [24].

We start by defining a walk between two individuals s and t on a multiplex network as a sequence of nodes, following intralayer and/or interlayer edges, which starts at node s “in any layer” and finishes in node t “in any layer.” Note that in this definition we do not care about the layer, just the node. The reasoning behind this lack of discrimination is that, in the multiplex structure, the different node replicas in the different layers correspond to the same individual (social networks) or location (transportation networks), thus it is only important to know if the packet has arrived, but not in which layer. **Figure 1** shows an example of a walk between two nodes in a multilayer network where non-trivial effects can be observed because of the presence of interlayer connections that affect the navigation through the networked system [17].

Given the definition of a walk in the multiplex topology, the effective betweenness of a node i in layer α , $B_{i\alpha}$, can be directly obtained as the fraction of walks that pass through node i in layer α for every possible origin-destination pair (s, t) . In some cases it might be convenient to obtain the betweenness of node i irrespective of the layer. In this case, the betweenness can be obtained just by accumulation of the individual contributions of each layer where i is represented, $B_i = \sum_{\alpha} B_{i\alpha}$.

For the specific case of the *shortest path betweenness*, every walk is restricted to be the path with minimum distance that

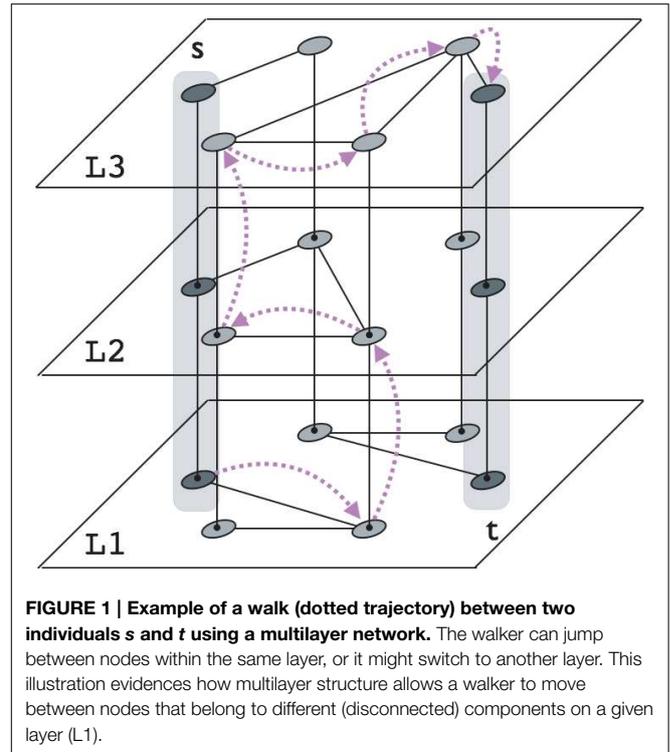


FIGURE 1 | Example of a walk (dotted trajectory) between two individuals s and t using a multilayer network. The walker can jump between nodes within the same layer, or it might switch to another layer. This illustration evidences how multilayer structure allows a walker to move between nodes that belong to different (disconnected) components on a given layer (L1).

starts from the source node s in any layer, and reaches the destination node t in any layer. The distance function may take into account the weights of the edges the path traverses. In this work, without loss of generality, we assume the edges’ weights are unitary and define the distance as the number of traversed edges in the path. The shortest path between two locations may be degenerated and consequently the set of shortest paths may contain paths using a single layer (classical shortest paths) and paths which change layer (pure multiplex paths). For an accurate computation of the shortest path betweenness special care must be taken with the path degeneration. A good and efficient algorithm can be found in Solé-Ribalta et al. [25].

Equivalently to the shortest path betweenness, the *random walk betweenness* depends on the particular definition of the network traversal procedure. In this case, a random walk is defined as a walk in which, at each time step, the next visited node is chosen with uniform random probability among the neighbors of the last visited node. The random walk betweenness is usually computed considering a transition matrix obtained from the adjacency matrix of the network. For a detailed description of random walks in a multiplex network, see [17]. In this document we will use the classic random walk definition. For the random walk betweenness the walk degeneracy is enormous and consequently is impractical to compute the betweenness accounting for all the possible individual random walks. Fortunately, the random walk betweenness can be efficiently computed using matrix inversion and absorbing random walks [26].

2.3. Community Structure in Multiplex

Networks representative of complex systems are characterized by having community structure, meaning the presence of dense groups of nodes with sparse connections between them [27]. It is known that dynamical processes running on top of such networks have a big dependency on community structure, which affects the process either by fostering or hampering it [28–30]. As evidenced in several works [7, 31], when the different layers of a multiplex network exhibit community structure, the influence on the overall system is not trivial to determine.

Here, to uncover the basic effects of communities in information flow process, we propose a simple setting with imposed community structure where communities fully overlap between layers and the degree of each node of the network is kept constant. Each multiplex network consists of two layers, and each layer has 256 nodes distributed in four communities (64 nodes per community) [32]. The links are generated in such a way that the density of links inside the communities is always higher than the density between them. The networks are generated independently for each layer, resulting in a two-layer multiplex network with different community structure in each layer.

For the experiments, we consider 12 different multiplex community structures and 300 different realizations for each. For all of them, we fix the bottom layer (L1) to $k_{in} = 31$ and $k_{out} = 1$ (i.e., 31 edges inside the community and 1 link outside, per node), which displays strong and clear communities, and we vary the community structure of the top layer (L2), which ranges from the previous strong block structure to a more diluted one ($k_{in} = 20$ and $k_{out} = 12$) where the communities are almost imperceptible. We quantify the strength of the community structure of the L2 layer using a mixing parameter defined as $\mu = k_{out}/\langle k \rangle$. **Figure 2** depicts three examples of such generated networks.

3. Results

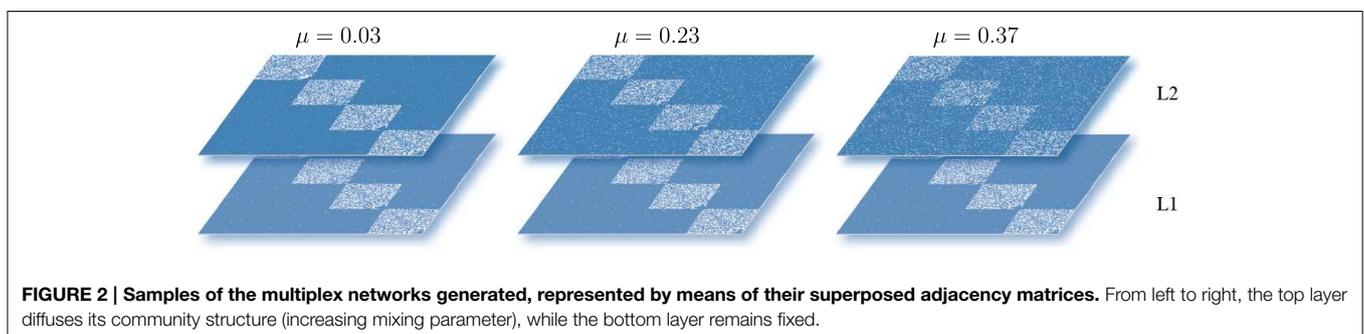
To evaluate the influence of the multiplex networks with community structure in information transfer we assess several aspects of the information transfer dynamics, namely the shortest paths distribution, the packets ingoing rate of each node and the critical injection rate of the network.

Figure 3 shows the obtained distribution of shortest paths in the different layers of the multiplex. In the case of having equivalent community structure in both layers (leftmost points in the plot), the multiplex structure provides a very good load balance where the same fraction of paths traverse using layer 1 and 2. In general, we can conclude that the effect of the multiplex is negligible for the overall system behavior since only a very small fraction of paths (0.5%) makes use of the full multiplex structure. In fact, paths using both layers of the multiplex are only used in the case where the origin and destination are in different communities. As we increase the mixing parameter of the second layer, its community structure dilutes, enhancing the communication between communities but slightly hindering the transfer of information internally. This effect is evident in **Figure 3**, which shows a large increase of intercommunity trips in the second layer and a small increase of intracommunity paths in the first layer. At the same time, the improvement of intercommunity paths in the more diffuse layer yields a disappearance of the (already small number of) shortest paths using both layers.

To assess the microscopic behavior of the system we show how the ingoing rate of packets to each node varies with respect to the mixing parameter. We compute the ingoing rate of each node of the multiplex structure as

$$\hat{\sigma}_{i\alpha} = \frac{L B_{i\alpha}}{N - 1}. \quad (3)$$

Results are shown in **Figure 4**. For the shortest path routing strategy (subplot A) we observe a clear distinction between the behavior of nodes in layers 1 and 2. As can be seen in **Figure 3**, the main effect on the increasing of the mixing parameter is clearly a migration of the shortest paths from layer 1 to layer 2, i.e., paths that traversed layer 1 now find a more efficient route through layer 2, which has a more diluted community structure. This migration of paths should increase the ingoing rate of nodes in layer 2 similarly to the observed decrease of ingoing packets of layer 1. This is the situation for small mixing parameters, but increasing the mixing parameter means also an increase in the efficiency of layer 2 routing packets between nodes in different communities, which in turn substantially reduces the overall node betweenness provoking an interesting tradeoff that will prescribe the final efficiency of the full system.



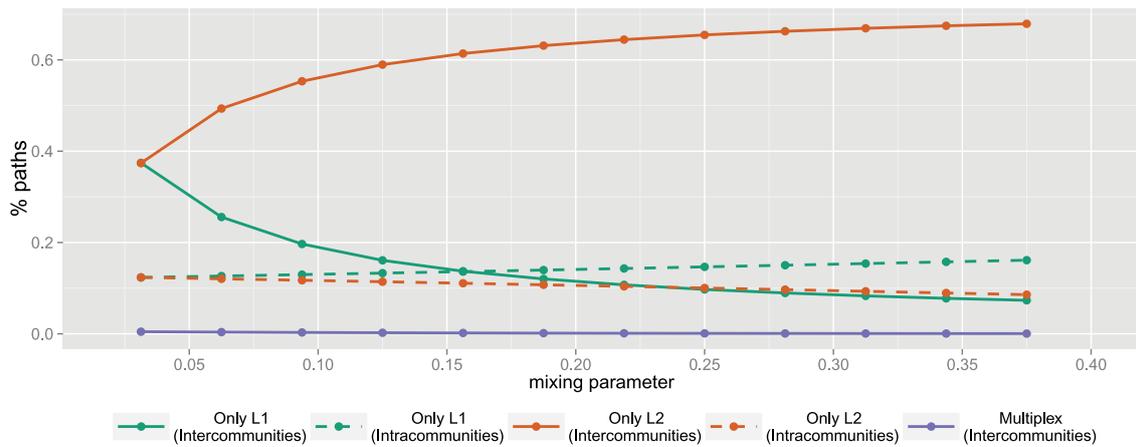


FIGURE 3 | Shortest paths distribution in a multiplex network with community structure as a function of the mixing parameter.

The plot shows the fraction of shortest paths that traverse the network using only layer 1 (with fixed topology), using only layer 2 (with varying topology) and using the full multiplex structure. The

horizontal axis corresponds to the mixing parameter. For the paths that only use a single layer, we divide the contribution between paths where the source and target nodes are within the same community and in different communities. There are no intracommunity paths that use both layers.

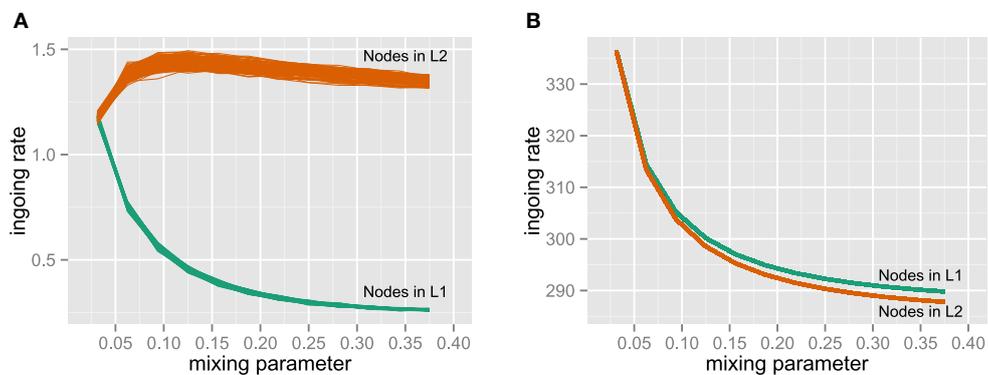


FIGURE 4 | Ingoing rate of each node in the network for different mixing parameter values, and two routing strategies: shortest paths (A) and random walks (B). The different colors indicate the layer to which the node belongs to.

These two opposed effects (migration of shortest paths and reduction of node betweenness) have a huge impact on the ingoing rate of nodes in layer 2, which experience a constant decrease after the maximum ingoing rate is reached. For the random routing strategy the scenario is completely different. The increase of the mixing parameter has an equivalent impact in both layers, which experience an important decrease of the ingoing rate.

Figures 5, 6 show the effect of the community structure on the critical injection rate ρ_c . For the shortest path routing strategy (Figure 5) the critical injection rate of the multiplex network reaches its minimum value around $\mu = 0.1$. This minimum indicates that there is a worst-case scenario for which the multiplex topology is less efficient than the individual layers. On the other side, the behavior of the critical injection rate of layer 2 is monotonically increasing. This situation is expected since a less clear community structure

leads to a reduction of the average shortest path, which in turn is positively correlated with a decrease of the node betweenness.

In general, if we compare the values of ρ_c for the multiplex network and the separate layers L1 and L2, we clearly observe three possible situations: (i) the multiplex is more resilient to congestion (efficient) than the individual layers. This situation arises when both layers have a similar community structure. (ii) The multiplex is less efficient than any of the layers. This setup corresponds to the minimum resilience of the multiplex structure. And (iii), the multiplex efficiency achieves a medium value which is a trade-off between the resilience of both layers. In a real setup, this situation would mean that joining those two layers in a multiplex improves the resilience observed in one layer at the cost of deteriorating the resilience observed in the second layer. However, as we can observe in the plot, the reduction of the efficiency is larger than the average of the

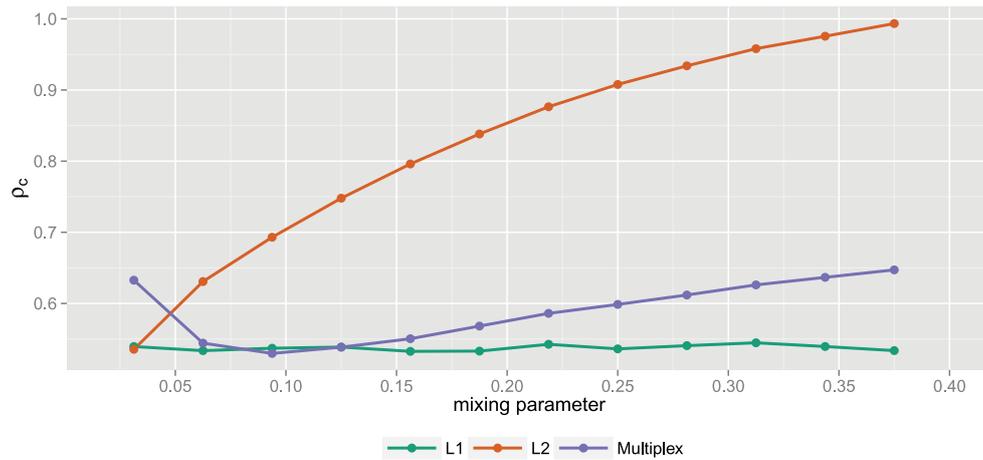


FIGURE 5 | Comparison of the critical injection rate of layer 1, layer 2 and the multiplex for different mixing parameter values and shortest path routing strategy. The values of the critical injection rate for single layer network and multiplex networks are computed using Equations (1) and (2) respectively.

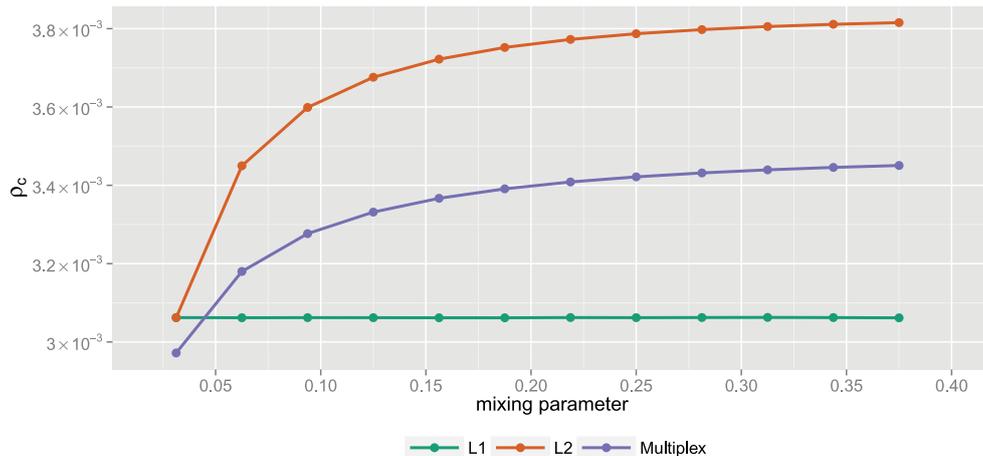


FIGURE 6 | Comparison of the critical injection rate of layer 1, layer 2 and the multiplex for different mixing parameter values and random walk routing strategy. The values of the critical injection rate for single layer network and multiplex networks are computed using Equations (1) and (2) respectively.

efficiency of both layers and consequently the coupling of layers is inefficient.

With respect to the random walk routing strategy (Figure 6), the situation is completely different. For similar community structures the multiplex worsens the efficiency, because the paths get trapped within the communities. For different community structures, in general, we obtain an efficiency that corresponds to the average efficiency of both layers.

4. Discussion

We have shown that packet information flow can be compromised when community structure is considered in some layer of the multiplex network structure. Since

community structure implies the presence of topological bottlenecks, the information flows migrate to those layers where these constraints are relaxed (diluted communities). We have shown that community structures produce a non-trivial effect in the transfer of information and in the resilience to information flow congestion, that here defines the efficiency of the structure. Essentially, the better defined the communities, the more affected the packet transportation. Information tries to avoid bottlenecks and packets migrate toward the layer where the community structure is diluted, because it is more efficient, but as a direct consequence of this migration the most efficient layer becomes overloaded. This nonlinear relation makes the problem of assessing the performance of the multiplex structure particularly challenging.

Using the analytical approach presented, we are able to determine for any multiplex topology what is the onset of congestion in the information flow and how it compares with the onset of the individual layers. We have also provided results, for very specific scenarios, of shortest path and random walk routing strategies respectively. The results show that the shortest path approach heavily depend on the particular sharpness of the prescribed communities. This work provides the starting point for the discrete flow analysis of more complicated scenarios of community structure in multiplex networks.

References

- Guimera R, Arenas A, Diaz-Guilera A. Communication and optimal hierarchical networks. *Phys A Stat Mech Appl.* (2001) **299**:247–52. doi: 10.1016/S0378-4371(01)00302-8
- Arenas A, Danon L, Diaz-Guilera A, Guimerà R. Local search with congestion in complex communication networks. In: Bubak M, van Albada GD, Sloot PMA, Dongarra J, editors. *Computational Science-ICCS 2004*. Kraków: Springer (2004). p. 1078–85. doi: 10.1007/978-3-540-24688-6_139
- Guimerà R, Diaz-Guilera A, Vega-Redondo F, Cabrales A, Arenas A. Optimal network topologies for local search with congestion. *Phys Rev Lett.* (2002) **89**:248701. doi: 10.1103/PhysRevLett.89.248701
- Kivela M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA. Multilayer networks. *J Comp Netw.* (2014) **2**:203–71. doi: 10.1093/comnet/cnu016
- Boccaletti S, Bianconi G, Criado R, del Genio CI, Gómez-Gardeñes J, Romance M, et al. The structure and dynamics of multilayer networks. *Phys Rep.* (2014) **544**:1–122. doi: 10.1016/j.physrep.2014.07.001
- Gauvin L, Panisson A, Cattuto C. Detecting the community structure and activity patterns of temporal networks: a non-negative tensor factorization approach. *PLoS ONE* (2014) **9**:e86028. doi: 10.1371/journal.pone.0086028
- Mucha PJ, Richardson T, Macon K, Porter MA, Onnela JP. Community structure in time-dependent, multiscale, and multiplex networks. *Science* (2010) **328**:876–8. doi: 10.1126/science.1184819
- Godoy-Lorite A, Guimera R, Sales-Pardo M. Long-term evolution of technological networks: statistical regularities, predictability and stability of social behaviors. (2015). arXiv:1506.01516
- Wellman B, Berkowitz SD. *Social Structures: A Network Approach*. Vol. 2. CUP Archive (1988).
- Brummitt CD, Lee KM, Goh KI. Multiplexity-facilitated cascades in networks. *Phys Rev E* (2012) **85**:045102. doi: 10.1103/PhysRevE.85.045102
- Yağan O, Gligor V. Analysis of complex contagions in random multiplex networks. *Phys Rev E* (2012) **86**:036103. doi: 10.1103/PhysRevE.86.036103
- Lambiotte R, Delvenne JC, Barahona M. Random walks, markov processes and the multiscale modular organization of complex networks. *Netw Sci Eng IEEE Trans.* (2014) **1**:76–90. doi: 10.1109/TNSE.2015.2391998
- Lambiotte R, Salnikov V, Rosvall M. Effect of memory on the dynamics of random walks on networks. *J Comp Netw.* (2015) **3**:177–88. doi: 10.1093/comnet/cnu017
- Bagnoli F, Massaro E. Risk perception and epidemic spreading in multiplex networks. In: *ISCS 2014: Interdisciplinary Symposium on Complex Systems*. Florence: Springer (2015). p. 319–32. doi: 10.1007/978-3-319-10759-2_33
- Zhao D, Li L, Peng H, Luo Q, Yang Y. Multiple routes transmitted epidemics on multiplex networks. *Phys Lett A* (2014) **378**:770–6. doi: 10.1016/j.physleta.2014.01.014
- Gómez S, Diaz-Guilera A, Gómez-Gardeñes J, Pérez-Vicente CJ, Moreno Y, Arenas A. Diffusion Dynamics on Multiplex Networks. *Phys Rev Lett.* (2013) **110**:028701. doi: 10.1103/PhysRevLett.110.028701
- De Domenico M, Solé-Ribalta A, Gómez S, Arenas A. Navigability of interconnected networks under random failures. *Proc Natl Acad Sci USA.* (2014) **111**:8351–6. doi: 10.1073/pnas.1318469111
- Solé-Ribalta A, De Domenico M, Kouvaris NE, Diaz-Guilera A, Gómez S, Arenas A. Spectral properties of the Laplacian of multiplex networks. *Phys Rev E* (2013) **88**:032807. doi: 10.1103/PhysRevE.88.032807
- Newman M. *Networks: An Introduction*. New York, NY, USA: Oxford University Press, Inc. (2010). doi: 10.1093/acprof:oso/9780199206650.001.0001
- Halu A, Mondragón RJ, Panzarasa P, Bianconi G. Multiplex PageRank. *PLoS ONE* (2013) **8**:e78293. doi: 10.1371/journal.pone.0078293
- Battiston F, Nicosia V, Latora V. Structural measures for multiplex networks. *Phys Rev E* (2014) **89**:032804. doi: 10.1103/PhysRevE.89.032804
- Sola L, Romance M, Criado R, Flores J, del Amo AG, Boccaletti S. Eigenvector centrality of nodes in multiplex networks. *Chaos* (2013) **3**:033131. doi: 10.1063/1.4818544
- De Domenico M, Solé-Ribalta A, Omodei E, Gómez S, Arenas A. Ranking in interconnected multilayer networks reveals versatile nodes. *Nat Commun.* (2015) **6**:6868. doi: 10.1038/ncomms7868
- De Domenico M, Solé-Ribalta A, Cozzo E, Kivela M, Moreno Y, Porter MA, et al. Mathematical formulation of multilayer networks. *Phys Rev X* (2013) **3**:041022. doi: 10.1103/PhysRevX.3.041022
- Solé-Ribalta A, De Domenico M, Gómez S, Arenas A. Centrality rankings in multiplex networks. In: *ACM Web Science Conference, WebSci '14*, Bloomington, IN (2014). p. 149–55. doi: 10.1145/2615569.2615687
- Solé-Ribalta A, De Domenico M, Gómez S, Arenas A. Random walk centrality in interconnected multilayer networks. (2015). arXiv:1506.07165
- Fortunato S. Community detection in graphs. *Phys Rep.* (2010) **486**:75–174. doi: 10.1016/j.physrep.2009.11.002
- Liu Z, Hu B. Epidemic spreading in community networks. *Eur Lett.* (2005) **72**:315. doi: 10.1209/epl/i2004-10550-5
- Arenas A, Diaz-Guilera A, Pérez-Vicente CJ. Synchronization Reveals Topological Scales in Complex Networks. *Phys Rev Lett.* (2006) **96**:114102. doi: 10.1103/PhysRevLett.96.114102
- Arenas A, Diaz-Guilera A, Prez-Vicente CJ. Synchronization processes in complex networks. *Phys D Nonlin Phenom.* (2006) **224**:27–34. doi: 10.1016/j.physd.2006.09.029
- Bennett L, Kittas A, Muirhead G, Papageorgiou LG, Tsoka S. Detection of composite communities in multiplex biological networks. *Sci. Rep.* (2015) **5**:10345. doi: 10.1038/srep10345
- Newman MEJ, Girvan M. Finding and evaluating community structure in networks. *Phys Rev E* (2004) **69**:026113. doi: 10.1103/PhysRevE.69.026113

Author Contributions

All the authors contributed equally to the article.

Acknowledgments

This work has been supported by Ministerio de Economía y Competitividad (Grant FIS2012-38266) and European Commission FET-Proactive Projects PLEXMATH (Grant 317614). AA also acknowledges partial financial support from the ICREA Academia and the James S. McDonnell Foundation.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Solé-Ribalta, Granel, Gómez and Arenas. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Temporal pattern of online communication spike trains in spreading a scientific rumor: how often, who interacts with whom?

Ceyda Sanli* and Renaud Lambiotte

Complexity and Networks, naXys, Department of Mathematics, University of Namur, Namur, Belgium

OPEN ACCESS

Edited by:

Javier Borge-Holthoefer,
Qatar Computing Research Institute,
Qatar

Reviewed by:

Marco Alberto Javarone,
University of Cagliari, Italy
David Garcia,
ETH Zurich, Switzerland

*Correspondence:

Ceyda Sanli,
Complexity and Networks, naXys,
Department of Mathematics,
University of Namur, Reimpart de la
Vierge 8, Namur 5000, Belgium
cedaysan@gmail.com

Specialty section:

This article was submitted to
Interdisciplinary Physics,
a section of the journal
Frontiers in Physics

Received: 31 July 2015

Accepted: 07 September 2015

Published: 25 September 2015

Citation:

Sanli C and Lambiotte R (2015)
Temporal pattern of online
communication spike trains in
spreading a scientific rumor: how
often, who interacts with whom?
Front. Phys. 3:79.
doi: 10.3389/fphy.2015.00079

We study complex time series (spike trains) of online user communication while spreading messages about the discovery of the Higgs boson in Twitter. We focus on online social interactions among users such as retweet, mention, and reply, and construct different types of active (performing an action) and passive (receiving an action) spike trains for each user. The spike trains are analyzed by means of local variation, to quantify the temporal behavior of active and passive users, as a function of their activity and popularity. We show that the active spike trains are bursty, independently of their activation frequency. For passive spike trains, in contrast, the local variation of popular users presents uncorrelated (Poisson random) dynamics. We further characterize the correlations of the local variation in different interactions. We obtain high values of correlation, and thus consistent temporal behavior, between retweets and mentions, but only for popular users, indicating that creating online attention suggests an alignment in the dynamics of the two interactions.

Keywords: social dynamic behavior, twitter social network, time series analysis, communication types in twitter, classifying active and popular users, ranking activation and popularity

1. Introduction

In recent years, online social media (OSM) have become a major communication channel, allowing users to share information in their social and professional circles, to discover relevant information pre-filtered by other users, and to chat with their acquaintances. In addition to their practical use for individuals, OSM have the advantage of generating a rich data set on collective social dynamics, as social relations among individuals, temporal properties of their interactions, and their contents are automatically stored. The study of these digital footprints has led to the emergence of computational social science, allowing to quantify at large-scales our political ideas and preferences [1], to discover roles in social networks [2, 3], to predict our health [4] and personality [5], and to determine external effects on online behavior [6]. Importantly, in OSM, users are at the same time both actors and receivers and therefore the amplification of a trend originates from the interplay between influencing [7, 8] and being influenced [9–13].

A crucial aspect of OSM and more generally of human behavior is the underlying complex dynamics [14–17]. The time series of user activities, e.g., posting a tweet and replying to a message, are quite distinct from uncorrelated (Poisson random) dynamics in the presence of burstiness [18–20], temporal correlations [6, 21, 22], and non-stationarity of human daily rhythm [23, 24], which has significant implications. Diffusion on a temporal network cannot be

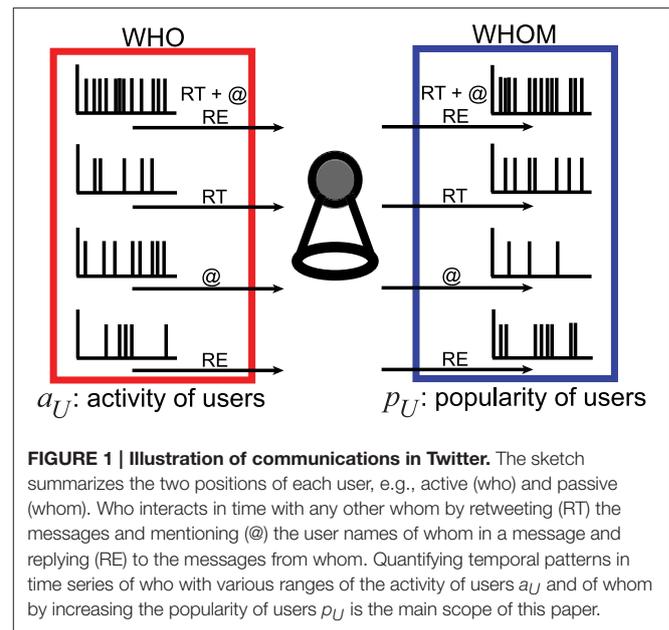
accurately described by models on static networks and consequently the process presents non-Markovian features with strong influence on the time required to explore the system [25, 26]. Furthermore, the dynamics drives a strong heterogeneity observed in user activity [27, 28] and user/content popularity [29–31]. Specifically, in Twitter, the heterogeneity in popularity has been observed and quantified in different ways by the size of retweet cascades, i.e., users re-transfer messages to their own followers with or without modifying them [32–36] or by the number of mentions of a user name, identified by the symbol @, in other people’s tweets [37].

In this paper, we focus on the dynamics of social interactions taking place when diffusing rumors about the discovery of the Higgs boson on July 2012 in Twitter [38]. Our main goal is to find connections between the statistical properties of user time series established on the same subject, e.g., the announcement of the discovery of the Higgs boson, and their activity and popularity. To this end, we analyze tweets including social interactions, such as retweets of a message (RT), mentions of a user name (@), and replies to a message (RE). For each type of the interactions, a user can either play an active, e.g., retweeting, or a passive, e.g., being retweeted, role. Therefore, we characterize each user by 8 time series: one active and one passive time series for each of the 3 types of interaction as well as for the aggregation of all interactions, as illustrated in **Figure 1**. Active time series are denoted as WHO and passive time series are defined by WHOM. We then investigate whether the statistical properties of each signal is a good predictor for the activity and popularity of a user.

The following sections are organized as follows. In Section 2, we describe the data set and provide basic statistical properties of who and whom time series. In Section 3, we introduce a technique dedicated to the analysis of non-stationary time series, so-called local variation, originally established for neuron spike trains [39–42] and recently applied to hashtag spike trains in Twitter [43, 44]. In Section 4, we search for statistical relations between local variation and measures of popularity of a user. Finally, Section 5 summarizes the key results and raises open questions.

2. Activity and Popularity of Users

Our aim is to examine the dynamics of user communications in Twitter. We investigate how frequently users talk to each other on a certain topic, e.g., the discovery of the Higgs boson, and identify how complex dynamic patterns of the communications evolve in time. To this end, we focus on the three different types of interaction between users, retweet (RT), mention (@), and reply (RE). Twitter users can adopt a tweet of someone and use it again in their own tweet by RT or contact to other users directly by typing user names in a message called @ or simply RE to any tweets, e.g., regular tweets, retweets, and tweets/retweets including @s. Typically, @s and REs are associated to personal interactions between users, whereas RTs are responsible for large-scale information diffusion in the social network and for the emergence of cascades. Here, we count all types of interaction as a part of complex information diffusion in the network.

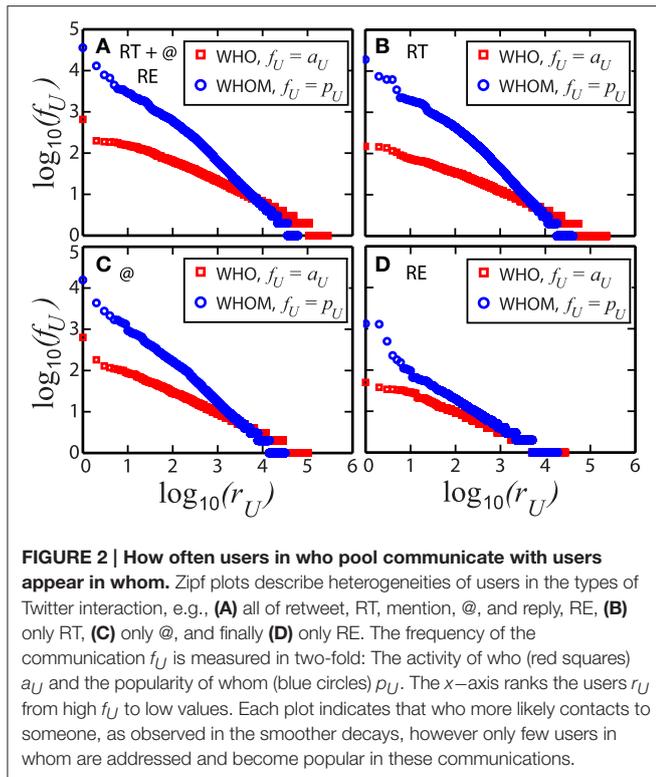


Interactions in Twitter are performed between at least two users (for instance, a user can mention several other users in a single tweet). Each action is directed and characterized by its timestamp. The users performing the action play active roles (who users), the users receiving their attention play a passive role (whom users), and each user can appear in both active and passive roles described in **Figure 1**. Therefore, we construct active and passive RT, @, and RE spike trains for each user.

2.1. Data Set

As a test bed, we consider the publicly available Higgs Twitter data set [38, 45], first collected to track the spread of the rumor on the discovery of the Higgs boson via RT, @, or RE. The data set is composed of tweets containing one of the following keywords or hashtags related to the discovery of the Higgs boson, “lhc,” “cern,” “boson,” and “higgs.” The start date is the 1st July 2012, 00:00 a.m. and the final date is the 7th July 2012, 11:59 p.m., which covers the announcement date of the discovery, the 4th July 2012, 08:00 a.m. All dates and timestamps in the data are converted to the Greenwich mean time. Detailed information on the data collection procedure and basic statistics can be found in Domenico et al. [38].

In total, the data is composed of 456,631 users (nodes) and 563,069 interactions (edges). Among those, we detect 354,930 RT, 171,237 @, and 36,902 RE, which shows that RT is more popular than the other communication channels. For RT interactions, we find 228,560 users join in who, in contrast, only 41,400 users appear in whom. These numbers are smaller for @, e.g., 102,802 who and 31,477 whom, and even smaller for RE, with 27,227 who and 18,578 whom. In each case, whom is much lower than who, as expected because a small number of users tend to attract a large fraction of attention in both friendship [46, 47] and online social [48–52] networks. This observation is confirmed in **Figure 2**, where we present Zipf plots associated to each



interaction, clearly showing a strong heterogeneity in the system. For who, the frequency of the user communication f_U ranks how active users are and measures the activity of users a_U , on the other hand, for whom, f_U quantifies how often the users or their tweets are addressed and so gives the popularity of users p_U .

3. Local Variation of Who and Whom

3.1. Communication Spike Trains

Evaluating each directed interaction (RT, @, and RE) of the users in the pool of who with any users in the whom class, as sketched in **Figure 1**, we extract salient temporal patterns of the user communication time series. We don't check whether the whom participates in the conversation in a later stage and only construct independent time series of the individual who and whom. The elements of the time series are the timestamps of the data [38, 45] providing us the exact time in second of the interaction and the user name or ID of the corresponding who and whom. Ordering the timestamps from the earliest to the latest, we generate spike trains carrying full story of the communication of each user. The resultant user communication spike trains are grouped in eight: For each who and whom, the spike trains of all interactions together (i) and the spike trains of filtered timestamps of RT (ii), @ (iii), and RE (iv).

3.2. Local Variation

A standard way of investigating the dynamics of human communication is to examine the statistics of the inter-event spike intervals such as its probability distribution [14], short-range memory coefficient and burstiness parameter [15] or

Fano factor. However, recent works have showed that further detail analysis is required to resolve temporal correlations [31, 32], bursts [19–22], and cascading [53] driven by circadian rhythm [23, 24], complex decision-making of individuals [3, 27, 54], and external factors [6] such as the announcement of discoveries, as considered in the current data [38].

To uncover the dynamics of the communication spike trains elaborately, we apply the local variation L_V originally defined to characterize non-stationary neuron spike trains [39–42] and very recently has been used to analyze hashtag spike trains [43, 44]. Unlike the memory coefficient and burstiness parameter [15], L_V provides a local temporal measurement, e.g., at τ_i of a successive time sequence of a spike train $\dots, \tau_{i-1}, \tau_i, \tau_{i+1}, \dots$, and so compares temporal variations with their local rates [41]

$$L_V = \frac{3}{N-2} \sum_{i=2}^{N-1} \left(\frac{(\tau_{i+1} - \tau_i) - (\tau_i - \tau_{i-1})}{(\tau_{i+1} - \tau_i) + (\tau_i - \tau_{i-1})} \right)^2 \quad (1)$$

where N is the total number of spikes. Equation (1) also takes the form [41]

$$L_V = \frac{3}{N-2} \sum_{i=2}^{N-1} \left(\frac{\Delta\tau_{i+1} - \Delta\tau_i}{\Delta\tau_{i+1} + \Delta\tau_i} \right)^2 \quad (2)$$

Here, $\Delta\tau_{i+1} = \tau_{i+1} - \tau_i$ quantifies the forward delays and $\Delta\tau_i = \tau_i - \tau_{i-1}$ represents the backward waiting times for an event at τ_i . Importantly, the denominator normalizes the quantity such as to account for local variations of the rate at which events take place. By definition, L_V takes values in the interval (0:3) [43]. It has been shown that helps at classifying dynamical patterns successfully [39, 40, 42–44]. Following the analysis of Gamma processes [39, 40, 43] conventional in neuron spike analysis [42], it is known that $L_V = 1$ for temporarily uncorrelated (Poisson random) irregular spike trains, and that higher values are associated to a burstiness of the spike trains. In contrast, smaller values indicate a higher regularity of the time series.

We now perform an analysis of L_V on the user communication spike trains. Equation (2) is performed through the spike trains with removing multiple spikes taking place within 1 s. Such events are rare and their impact on the value of L_V has been shown to be limited [43]. **Figure 3** describes the distribution of L_V , $P(L_V)$ of full spike trains all together with RT, @, and RE for the who (a, b) and whom (c, d). Grouping L_V based on the frequency f_U , e.g., the activity of the who a_U and the popularity of the whom p_U , we examine the temporal patterns of the trains in different classes of a_U and p_U . For the real data in (a, c), in **Figure 3A**, L_V is always larger than 1 in any values of a_U , suggesting that all users playing a role in who contact to the whom in bursty communications. However, in **Figure 3C**, we observe distinct behavior of the whom users and bursts present only for low p_U . By increasing p_U , $L_V \approx 1$ indicating that there is no temporal correlation among the who referring the whom and L_V is slightly smaller than 1 for the most popular users, indicating a tendency toward regularity in the time series, as also observed for the hashtag spike trains [43]. These observations

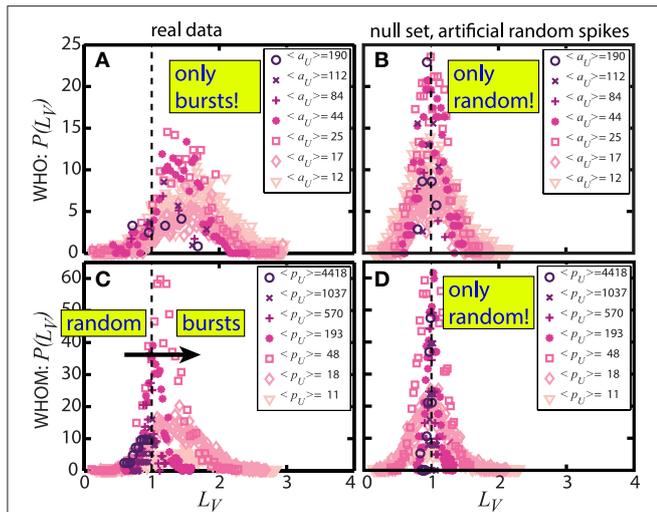


FIGURE 3 | Probability density function of the local variation L_V , $P(L_V)$ of who (A,B) and whom (C,D) users in various ranges of the two communication frequencies, e.g., a_U and p_U . (A,C) describe the results of the real data. When we only observe bursty communication patterns in who independently of the average user activity frequency (a_U) in (A), significant variations in L_V by increasing the average user popularity (p_U) are clear in (C). The results prove that popular users are addressed randomly in time and slightly more regular patterns observed in the most popular users. On the other hand, (B,D) present the statistics of artificially generated random spikes serving as a null model and all frequency ranges give the distributions around 1, as expected for temporarily uncorrelated signals.

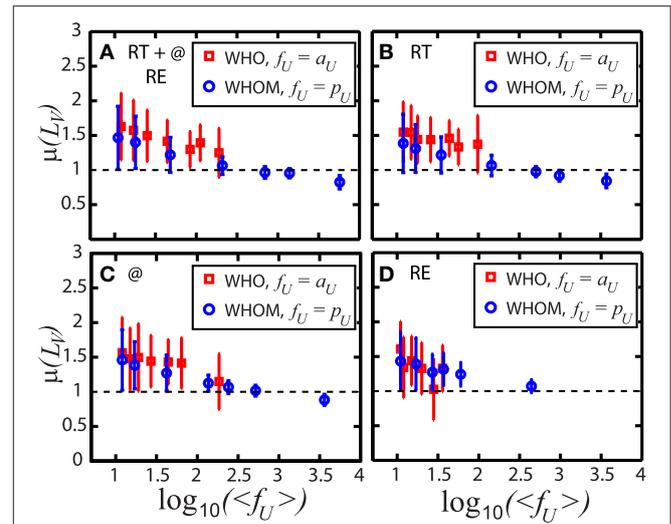


FIGURE 4 | Mean μ of the local variation L_V of the user communication spike trains vs. the logarithmic average frequency $\log_{10}(f_U)$. The results of who are represented by red squares and blue circles describe that of whom. Types of the interaction are investigated in detail: (A) All communications of retweet, RT, mention, @, and reply, RE. (B) Only RT. (C) Only @. (D) Only RE. Independent of the types of the interaction, the frequency of communication, e.g., the activity of users a_U and the popularity of users p_U , designs overall communication patterns. While low f_U gives bursty patterns with $L_V > 1$, moderate f_U indicates irregular uncorrelated (Poisson random) signals, e.g., $L_V \approx 1$. For all high f_U , $L_V < 1$ presenting the regularity of the communications. The error bars show the corresponding standard variations.

are significantly different for artificial spike trains constructed by randomly permuting the real full spike train and so expected to generate non-stationary Poisson processes. Therefore, all distributions are centered around 1 in this case, independently of a_U and p_U , as shown in Figures 3B,D. The randomization and obtaining a null set follow the same procedure explained in detail in Sanli and Lambiotte [43].

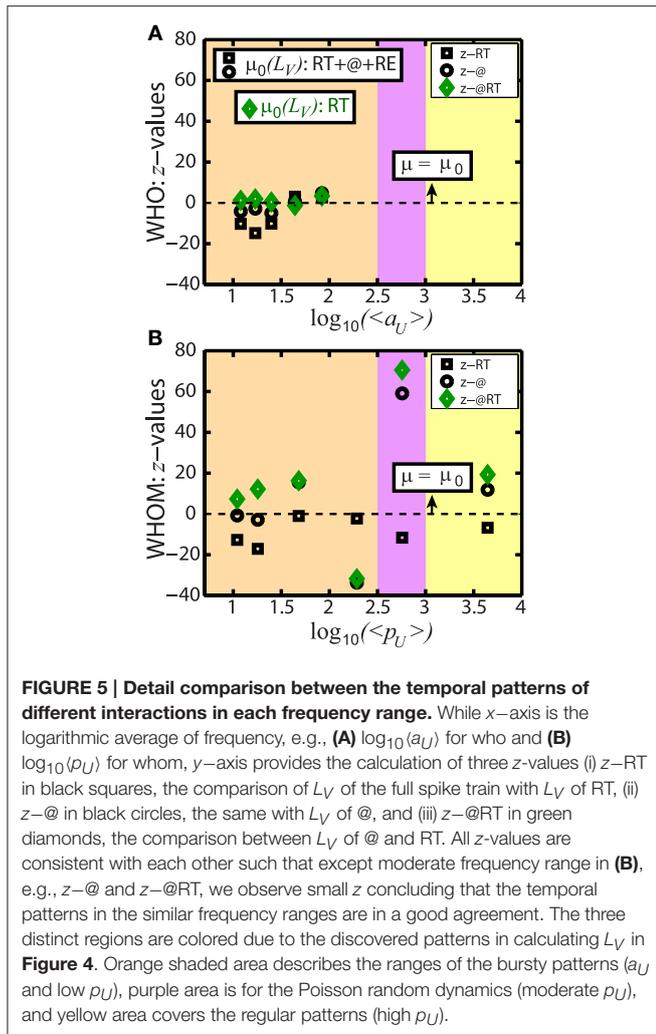
Even though Figure 3 represents $P(L_V)$ of full spike trains, i.e., all interactions together, $P(L_V)$ of individual RT, @, and RE communication spike trains describes very similar temporal behavior for both the who and whom. Figure 4 summarizes the detail of $P(L_V)$, the mean of L_V , $\mu(L_V)$ with the corresponding standard deviations $\sigma(L_V)$ as error bars, comparatively. The results highlight that to classify the communication temporal patterns neither the position of the users, whether active or passive, nor the types of the interaction, but the frequency of the communication f_U such as a_U and p_U plays a major role. All Figures 4A–D, we observe three regions: Bursts in low f_U , $\log_{10}(f_U) < 2.5$, irregular uncorrelated (Poisson random) dynamics in moderate and high f_U , $\log_{10}(f_U) \approx 2.5-3$, and regular patterns in very high f_U , $\log_{10}(f_U) > 3$. This conclusion supports the importance of frequency so time parameter overall human behavior [14, 16]. Applying standard linear fittings to the underlying data of Figure 4, composed of 5104 data points for whom, the understanding can be further proven. We observe the significant negative trend of L_V with increasing p_U , i.e., the slope is -0.32 .

We now perform a more thorough comparison in Figure 5, on the disparity of L_V in different frequency ranges. To this end, we calculate the standard z -values in two ways. First, to compare L_V of the full spike trains with L_V of only RT and also with L_V of only @ spike trains, L_V^{RT} and $L_V^@$, respectively, we introduce

$$z(f_U) = \frac{\mu(L_V^k) - \mu_0(L_V)}{\sigma(L_V^k) / \sqrt{f_U^k}} \quad (3)$$

Here, k in superscripts labels the interaction, e.g., either RT or @. Precisely, L_V^k is determined based on a filtered spike train composed of the user timestamps of either RT or @, as already used in (Figures 4B,C). In addition, μ^k is the mean of L_V^k , also presented in Figures 4B,C, and μ_0 is the mean L_V of the full spike train, given in Figure 4A.

In Figure 5, black squares show z -values of RT and black circles describe z -values of @. For who in Figure 5A where L_V only presents bursty patterns (orange shaded area) and low a_U , we have small z -values proving the agreement of the temporal patterns suggested by L_V in the same a_U . However, for whom in Figure 5B where we have rich values of p_U compared to the values of a_U , while z -values are small in bursty patterns (low p_U , orange area) as also in who and in regular patterns (high p_U , yellow area), larger z -@ value (the black circle) is calculated in uncorrelated Poisson dynamics (moderate p_U , purple area). The disagreement of L_V with large z -@ indicates that even though



$L_V \approx 1$ in this region the results of @ are quite sensitive in the same p_U , which is not observed in z -RT (the black square).

Furthermore, we repeat the analysis across communication channels by comparing temporal patterns of RT and @ as follows

$$z(f_U) = \frac{\mu(L_V^@) - \mu_0(L_V^{RT})}{\sigma(L_V^@)/\sqrt{f_U^@}} \quad (4)$$

The corresponding z -values, z -@RT are presented in green diamonds in Figure 5. Comparing to the previous z -RT and z -@, we now obtain even lower values for who (Figure 5A) showing a better agreement between RT and @ patterns. Moreover, we have a very similar trend for whom (Figure 5B) as before and so a large fluctuation is only observed in the purple area.

4. Correlation of L_V in User Communication Habits

In this final section, our interest turns into building new measures to quantify how the local variation L_V fluctuates inside

different classes of the frequency, f_U . What extend temporal communication habits of two independent users in the same f_U ranges agree with each other is the first question we address. Second, we examine whether the temporal patterns of the interactions are consistent for the same users and how the metric varies with increasing f_U .

We consider $r_{ij}^{kk'}(f_U)$, the Pearson correlation coefficient of L_V of two different users selected independently from the same f_U classes

$$r_{ij}^{kk'}(f_U) = \frac{\sum_{i,j=1, i \neq j}^{N_U} [L_{V_i}^k - \mu(L_{V_i}^k)][L_{V_j}^{k'} - \mu(L_{V_j}^{k'})]}{\sigma(L_{V_i}^k)\sigma(L_{V_j}^{k'})} \quad (5)$$

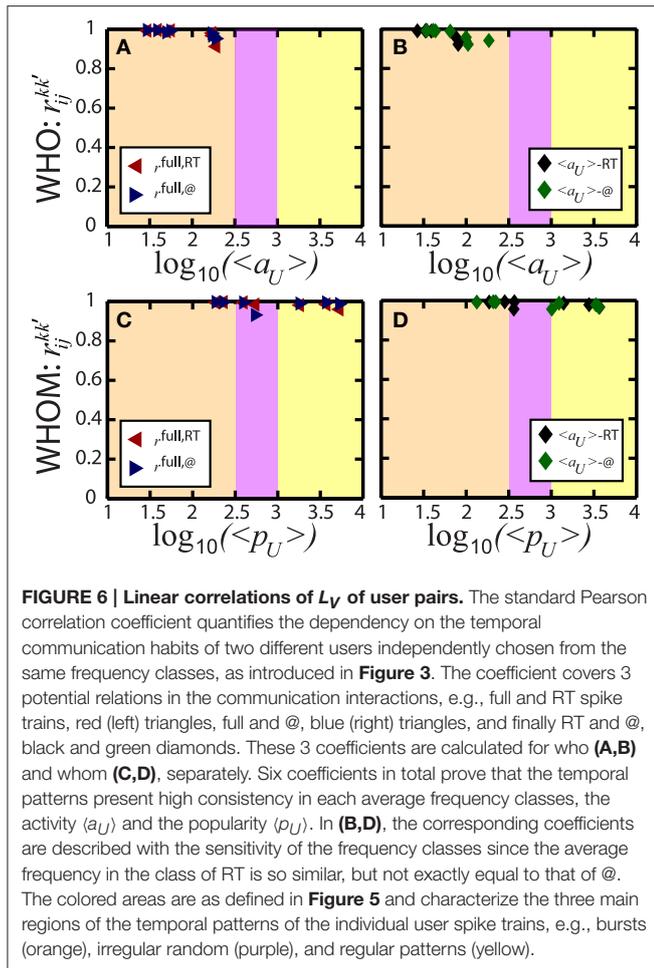
where $\sigma(L_{V_i}^k) = \sqrt{\sum_{i=1}^{N_U} [L_{V_i}^k - \mu(L_{V_i}^k)]^2}$. Here, L_{V_i} and L_{V_j} are the local variations of user i and j , respectively, μ 's are the corresponding mean values, and N_U is the total number of users. Moreover, k and k' represent all permutations among the full, RT, and @ spike trains. Furthermore, $r_{ij}^{kk'}(f_U)$ is evaluated for who and whom, separately. Therefore, i and j are different users, but from the same (who/whom) pool and in the same frequency classes of a_U and p_U , as grouped in Figure 3. Note that before performing Equation (5), the corresponding L_V 's in the same f_U class are ordered from the highest to the smallest (or vice versa) not to deform $r_{ij}^{kk'}(f_U)$ artificially due to the random selection.

Figure 6 presents the results of $r_{ij}^{kk'}(f_U)$ for who in (a, b) and whom in (c, d). Similar to z -values performed in the previous Section, we suggest three correlation coefficients: Red (left) triangles describe $r_{ij}^{\text{full,RT}}$, blue (right) triangles are for $r_{ij}^{\text{full,@}}$, and black and green diamonds show the values of $r_{ij}^{\text{RT,@}}$. The average frequency of the users $\langle f_U \rangle$ in the same class is similar but not equal and that is why Figures 6B,D are plotted with respect to both the mean frequencies of RT and @, e.g., the average activity $\langle a_U \rangle$ and popularity $\langle p_U \rangle$ of RT and @. All correlations are above 0.85 proving the high dependency of the communication patterns of the users in the same $\langle f_U \rangle$, independently of the types of the interaction.

We now consider Equation (5) with imposing the same user and repeat the procedure above for the correlation coefficient

$$r_i^{kk'}(f_U) = \frac{\sum_i^{N_U} [L_{V_i}^k - \mu(L_{V_i}^k)][L_{V_i}^{k'} - \mu(L_{V_i}^{k'})]}{\sigma(L_{V_i}^k)\sigma(L_{V_i}^{k'})} \quad (6)$$

Figure 7 summarizes the results of Equation (6). While Figures 7A,C are in parallel with that of Figure 6 with slightly lower correlations for @ (blue right triangles), distinct behavior is observed in Figures 7B,D. Low correlations in Figure 7B indicate that the same who users present different temporal behavior in RT and @. On the other hand, Figure 7D shows an interesting temporal habit of whom users. Having no remarkable dependency captured in low popular users, we show that the correlation increases with $\langle p_U \rangle$ describing that the popular users are addressed in RT and @ in a temporarily similar procedure.



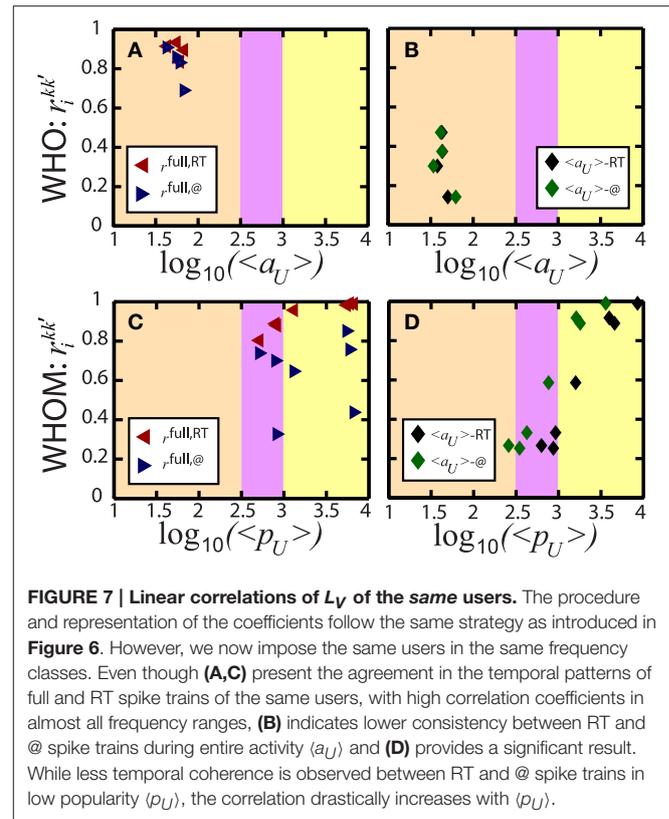
4.1. Nomenclature

- OSM: Online Social Media,
- @: Mention a user name in a tweet message,
- RE: Reply to a tweet or retweet message,
- RT: Retweet, share a message of other users in her/his own tweet blog,
- WHO: Twitter users starting an interaction via @ or RE or RT with any other users,
- WHOM: Twitter users addressed by who such that their message is retweeted or user name is mentioned in a message by who or they get a reply from who.

Any relation between who and whom such as the following-follower is not imposed.

5. Discussion

In this paper, our interest is to quantify online user communication in Twitter. To reduce the complexity in the communication, the data studied here consider only a unique subject which users talk about, that is the discovery of the Higgs



boson on July 4, 2012 within a restricted time window, e.g., 6 days [38]. The main aim is to extract salient temporal patterns of communication in various types of interaction observed in Twitter such as retweet (RT), mention (@), and reply (RE). Adopting the technique so-called local variation L_V originally introduced for neuron spike trains [39–42] and recently has applied to hashtag spike trains in Twitter [43, 44], we perform detailed analysis on user communication spike trains. Showing strong influences of the frequency of the hashtag spike trains on the resultant temporal patterns in the earlier work [43, 44], in parallel we here examine the differences in the patterns induced by the frequency of the user communication spike trains, f_U .

We investigate user communication spike trains in two categories, the first set of users are the active ones, who users, and the other set is composed of the passive users, whom users, in the communication, and each user can appear in both pools. For who, f_U simply gives what extend users contact to whom and so it is the activity of who, a_U . On the other hand, for whom, the generated spike trains present how often who refers the messages or the user names of whom and therefore, f_U is the popularity of whom, p_U . Providing comparative statistics on L_V of who and whom with increasing a_U and p_U , respectively, we observe quite distinct temporal behavior of online users. First, we observe an asymmetry between active and passive interactions, as only the former give rise to hubs, with few users attracting a large share of the attention. Moreover, who constantly presents bursts, $L_V > 1$ for all values of a_U , whereas whom demonstrates various dynamic behavior patterns, depending on

the popularity: The least popular users with low p_U experience bursty time series, popular users with moderate and high p_U are contacted by temporarily uncorrelated who users and so show Poisson random spike trains $L_V \approx 1$, and the most popular users with the maximum p_U are referred regularly in time, e.g., $L_V < 1$.

These scenarios are independent of both the position of users, e.g., who or whom, and the preferred interactions, e.g., whether RT or @, suggesting that the frequency of the communication dominates to design social dynamic behavior. This conclusion is also supported by high correlation coefficients of L_V on the user pairs in the same frequency classes. Furthermore, the linear correlation of L_V on the same users reveals interesting patterns. There, we observe that only popular users have similar dynamic behavior in both RT and @, which confirms that both metrics are complementary to characterize the influence of users.

The analysis could be improved by integrating the communication spike trains with the following-follower relation in Twitter, and focusing on the who and whom trains of connected users. An important concern is the limited time period of the data which the collection started 3 days before the announcement of the discovery and continued until 3 days after this date. Yet, it has been shown that the dynamics of the communication is drastically different before/after and during the announcement [38], and this variation could be investigated in our analysis. Our study shares the similar aims of the other research on online user behavior and the influence of the frequency in online platforms such as Flickr, Delicious and StumbleUpon, which user profiles have been included in the

analysis [47]. This understanding could be also applied to our analogy with considering further details in the data.

5.1. Data Sharing

The full data studied in this paper has open access [38, 45].

Author Contributions

Conceived and designed the experiments: CS. Performed the experiments: CS. Analyzed the data: CS. Contributed reagents/materials/analysis tools: RL, CS. Wrote the paper: CS, RL.

Funding

The EU 7th Framework OptimizR Project: 48909A2 CE OPTIMIZR (Grant holder: RL, Funding receiver: CS—<http://optimizr.eu/>) and FNRS MIS F4527.12 48888F3 (Grant holder: RL, Funding receiver: CS—<http://www.fnrs.be/>).

Acknowledgments

CS acknowledges supports from the European Union 7th Framework OptimizR Project and FNRS (le Fonds de la Recherche Scientifique, Wallonie, Belgium). This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office.

References

- Abisheva A, Garimella VRK, Garcia D, Weber I. Who watches (and shares) what on youtube? and when?: using twitter to understand youtube viewership. In: *Proceedings of the 7th ACM International Conference on Web Search and Data Mining, WSDM '14*. New York, NY: ACM (2014). pp. 593–602.
- Wu S, Hofman JM, Mason WA, Watts DJ. Who says what to whom on twitter. In: *Proceedings of the 20th International Conference on World Wide Web, WWW '11*. New York, NY: ACM (2011). pp. 705–14.
- Xuan Q, Gharehyazie M, Devanbu PT, Filkov V. Measuring the effect of social communications on individual working rhythms: a case study of open source software. In: *Social Informatics (SocialInformatics), 2012 International Conference on*. Washington, DC (2012). pp. 78–5.
- Choudhury MD, Gamon, Counts, Horvitz. Predicting depression via social media. In: Kiciman E, Ellison NB, Hogan B, Resnick P, Soboroff I, editors. *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*. Ann Arbor, MI: Association for the Advancement of Artificial Intelligence (AAAI) (2013). pp. 128–37.
- Lambiotte R, Kosinski M. Tracking the digital footprints of personality. *Proc IEEE* (2014) **102**:1934–9. doi: 10.1109/JPROC.2014.2359054
- Mathiesen J, Angheluta L, Ahlgren PTH, Jensen MH. Excitable human dynamics driven by extrinsic events in massive communities. *Proc Natl Acad Sci USA*. (2013) **110**:17259–62. doi: 10.1073/pnas.1304179110
- Bakshy E, Hofman JM, Mason WA, Watts DJ. Everyone's an influencer: quantifying influence on twitter. In: *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, WSDM '11*. New York, NY: ACM (2011). pp. 65–74.
- Gonzalez-Bailon S, Borge-Holthoefer J, Moreno Y. Broadcasters and hidden influentials in online protest diffusion. *Am Behav Sci*. (2013) **57**:943–65. doi: 10.1177/0002764213479371
- Bakshy E, Karrer B, Adamic LA. Social influence and the diffusion of user-created content. In: *Proceedings of the 10th ACM Conference on Electronic Commerce, EC '09*. New York, NY: ACM (2009). pp. 325–34.
- Lin YR, Bagrow JP, Lazer D. More voices than ever? quantifying media bias in networks. In: Nicolov N, Shanahan JG, editors. *Proceedings of the Fifth International AAAI Conference on Weblogs and Social Media*. Barcelona: Association for the Advancement of Artificial Intelligence (AAAI) (2011). pp. 193–200.
- Kwon KH, Stefanone MA, Barnett GA. Social network influence on online behavioral choices: exploring group formation on social network sites. *Am Behav Sci*. (2014) **58**:1345–60. doi: 10.1177/0002764214527092
- Zhang J, Tang J, Li J, Liu Y, Xing C. Who influenced you? predicting retweet via social influence locality. *ACM Trans Knowl Discov Data* (2015) **9**:25:1–26. doi: 10.1145/2700398
- Lerman K, Yan X, Wu XZ. The majority illusion in social networks. *ArXiv e-prints* (2015) **1506.03022**:1–10.
- Barabási AL. The origin of bursts and heavy tails in human dynamics. *Nature* (2005) **435**:207–11. doi: 10.1038/nature03459
- Goh KI, Barabasi AL. Burstiness and memory in complex systems. *Europhys Lett*. (2008) **81**:48002. doi: 10.1209/0295-5075/81/48002
- Miritello G, Lara R, Moro E. Time allocation in social networks: correlation between social structure and human communication dynamics. In: Holme P, Saramaki J, editors. *Temporal Networks*. Berlin; Heidelberg: Springer (2013). pp. 175–90.

17. Formentin M, Lovison A, Maritan A, Zanzotto G. Hidden scaling patterns and universality in written communication. *Phys Rev E* (2014) **90**:012817. doi: 10.1103/PhysRevE.90.012817
18. Wang P, Lei T, Yeung CH, Wang BH. Heterogenous human dynamics in intra- and inter-day time scales. *Europhys Lett.* (2011) **94**:18005. doi: 10.1209/0295-5075/94/18005
19. Zhou T, Zhao ZD, Yang Z, Zhou C. Relative clock verifies endogenous bursts of human dynamics. *Europhys Lett.* (2012) **97**:18006. doi: 10.1209/0295-5075/97/18006
20. Colman ER, Greetham DV. Memory and burstiness in dynamic networks. *Phys Rev E* (2015) **92**:012817. doi: 10.1103/PhysRevE.92.012817
21. Karsai M, Kaski K, Barabási AL, Kertesz J. Universal features of correlated bursty behaviour. *Sci Rep.* (2012) **2**:397. doi: 10.1038/srep00397
22. Jo HH, Perotti JL, Kaski K, Kertesz J. Correlated bursts and the role of memory range. *Phys Rev E* (2015) **92**:022814. doi: 10.1103/PhysRevE.92.022814
23. Jo HH, Karsai M, Kertesz J, Kaski K. Circadian pattern and burstiness in mobile phone communication. *New J Phys.* (2012) **14**:013055. doi: 10.1088/1367-2630/14/1/013055
24. Kim J, Lee D, Kahng B. Microscopic modelling circadian and bursty pattern of human activities. *PLoS ONE* (2013) **8**:e58292. doi: 10.1371/journal.pone.0058292
25. Iribarren JL, Moro E. Impact of human activity patterns on the dynamics of information diffusion. *Phys Rev Lett.* (2009) **103**:038702. doi: 10.1103/PhysRevLett.103.038702
26. Lambiotte R, Tabourier L, Delvenne JC. Burstiness and spreading on temporal networks. *Eur Phys J B* (2013) **86**:320. doi: 10.1140/epjb/e2013-40456-9
27. Wang C, Huberman BA. How random are online social interactions? *Sci Rep.* (2012) **2**:633. doi: 10.1038/srep00633
28. Formentin M, Lovison A, Maritan A, Zanzotto G. Universal activity pattern in human interactive dynamics. *ArXiv e-prints* (2014) **1405.5726**:1–5.
29. Ferrara E, Interdonato R, Tagarelli A. Online popularity and topical interests through the lens of instagram. In: *Proceedings of the 25th ACM Conference on Hypertext and Social Media, HT '14*. New York, NY: ACM (2014). pp. 24–34.
30. Ratkiewicz J, Fortunato S, Flammini A, Menczer F, Vespignani A. Characterizing and modeling the dynamics of online popularity. *Phys Rev Lett.* (2010) **105**:158701. doi: 10.1103/PhysRevLett.105.158701
31. Coscia M. Competition and success in the meme pool: a case study on quickmeme.com. In: *International AAAI Conference on Weblogs and Social Media* (2013). pp. 100–9. Available online at: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM13/paper/view/5990>
32. Myers SA, Leskovec J. The bursty dynamics of the twitter information network. In: *Proceedings of the 23rd International Conference on World Wide Web, WWW '14*. New York, NY: ACM (2014). pp. 913–24.
33. ten Thij M, Ouboter T, Worm D, Litvak N, van den Berg H, Bhulai S. Modelling of trends in twitter using retweet graph dynamics. In: Bonato A, Graham FC, Pralat P, editors. *Algorithms and Models for the Web Graph, Vol. 8882, of Lecture Notes in Computer Science*. Springer International Publishing (2014). pp. 132–47. Available online at: http://dx.doi.org/10.1007/978-3-319-13123-8_11
34. Zhao Q, Erdogdu MA, He HY, Rajaraman A, Leskovec J. SEISMIC: a self-exciting point process model for predicting tweet popularity. In: Fayyad U, Cao L, Zhang C, editors. *KDD15*. Sydney: ACM (2015).
35. Boyd D, Golder S, Lotan G. Tweet, tweet, retweet: conversational aspects of retweeting on twitter. In: *System Sciences (HICSS), 2010 43rd Hawaii International Conference on*. Honolulu, HI (2010). pp. 1–10.
36. Zaman TR, Herbrich R, Gael JV, Stern D. Predicting information spreading in twitter. In: *Computational Social Science and the Wisdom of Crowds Workshop (Colocated with NIPS 2010)* (2010). Available online at: <http://research.microsoft.com/apps/pubs/default.aspx?id=141866>
37. Greetham DV, Ward. Conversations on twitter: structure, pace, balance. In: Kanawati R, Pensa RG, Rouveiro C, editors. *Proceedings of the 2nd International Workshop on Dynamic Networks and Knowledge Discovery. CEUR Workshop Proceedings*. Nancy (2014). pp. 73–84.
38. Domenico MD, Lima A, Mougél P, Musolesi M. The anatomy of a scientific rumor. *Sci Rep.* (2013) **3**:2980. doi: 10.1038/srep02980
39. Shinomoto S, Shima K, Tanji J. Differences in spiking patterns among cortical neurons. *Neural Comput.* (2003) **15**:2823–42. doi: 10.1162/089976603322518759
40. Miura K, Okada M, Amari S. Estimating spiking irregularities under changing environments. *Neural Comput.* (2006) **18**:2359–86. doi: 10.1162/neco.2006.18.10.2359
41. Omi T, Shinomoto S. Optimizing time histograms for non-poissonian spike trains. *Neural Comput.* (2011) **23**:3125–44. doi: 10.1162/NECO.a.00213
42. Shinomoto S, Kim H, Shimokawa T, Matsuno N, Funahashi S, Shima K, et al. Relating neuronal firing patterns to functional differentiation of cerebral cortex. *PLoS Comput Biol.* (2009) **5**:e1000433. doi: 10.1371/journal.pcbi.1000433
43. Sanli C, Lambiotte R. Local variation of hashtag spike trains and popularity in twitter. *PLoS ONE* (2015) **10**:e0131704. doi: 10.1371/journal.pone.0131704
44. Sanli C, Lambiotte R. Local variation of collective attention in hashtag spike trains. In: *Modeling and Mining Temporal Interactions: Papers from the 2015 ICWSM Workshop* (2015). pp. 8–12.
45. Leskovec J, Krevl A. *SNAP Datasets: Stanford Large Network Dataset Collection* (2014). Available online at: <http://snap.stanford.edu/data>
46. Saramaki J, Leicht EA, Lopez E, Roberts SGB, Reed-Tsochas F, Dunbar RIM. Persistence of social signatures in human communication. *Proc Natl Acad Sci USA.* (2014) **111**:942–7. doi: 10.1073/pnas.1308540110
47. Meo Pd, Ferrara E, Abel F, Aroyo L, Houben GJ. Analyzing user behavior cross social sharing environments. *ACM Trans Intell Syst Technol.* (2014) **5**:14:1–31.
48. Huberman BA, Romero DM, Wu F. Social networks that matter: twitter under the microscope. *First Monday* (2009) **14**:1–9. doi: 10.5210/fm.v14i1.2317
49. Gonçalves B, Perra N, Vespignani A. Modeling Users' activity on twitter networks: validation of dunbar's number. *PLoS ONE* (2011) **6**:e22656. doi: 10.1371/journal.pone.0022656
50. Weng L, Flammini A, Vespignani A, Menczer F. Competition among memes in a world with limited attention. *Sci Rep.* (2012) **2**:335. doi: 10.1038/srep00335
51. Gleeson JP, Ward JA, O'Sullivan KP, Lee WT. Competition-Induced criticality in a model of meme popularity. *Phys Rev Lett.* (2014) **112**:048701. doi: 10.1103/PhysRevLett.112.048701
52. Cetin U, Bingol HO. Attention competition with advertisement. *Phys Rev E* (2014) **90**:032801. doi: 10.1103/PhysRevE.90.032801
53. Borge-Holthoefer J, Banos RA, Gonzalez-Bailon S, Moreno Y. Cascading behaviour in complex socio-technical networks. *J Complex Netw.* (2013) **1**–22. doi: 10.1093/comnet/cnt006
54. Jo HH, Pan RK, Kaski K. Time-varying priority queuing models for human dynamics. *Phys Rev E* (2012) **85**:066101. doi: 10.1103/PhysRevE.85.066101

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Sanli and Lambiotte. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Characterizing interactions in online social networks during exceptional events

Elisa Omodei*, Manlio De Domenico and Alex Arenas

Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Tarragona, Spain

OPEN ACCESS

Edited by:

Taha Yasseri,
University of Oxford, UK

Reviewed by:

Benjamin Miranda Tabak,
Universidade Católica de Brasília,
Brazil
Sandro Meloni,
University of Zaragoza, Spain

*Correspondence:

Elisa Omodei,
Departament d'Enginyeria Informàtica
i Matemàtiques, Universitat Rovira i
Virgili, Avda. Paisos Catalans, 26,
43007 Tarragona, Spain
elisa.omodei@urv.cat

Specialty section:

This article was submitted to
Interdisciplinary Physics,
a section of the journal
Frontiers in Physics

Received: 30 June 2015

Accepted: 27 July 2015

Published: 11 August 2015

Citation:

Omodei E, De Domenico M and
Arenas A (2015) Characterizing
interactions in online social networks
during exceptional events.
Front. Phys. 3:59.
doi: 10.3389/fphy.2015.00059

Nowadays, millions of people interact on a daily basis on online social media like Facebook and Twitter, where they share and discuss information about a wide variety of topics. In this paper, we focus on a specific online social network, Twitter, and we analyze multiple datasets each one consisting of individuals' online activity before, during and after an exceptional event in terms of volume of the communications registered. We consider important events that occurred in different arenas that range from policy to culture or science. For each dataset, the users' online activities are modeled by a multilayer network in which each layer conveys a different kind of interaction, specifically: retweeting, mentioning and replying. This representation allows us to unveil that these distinct types of interaction produce networks with different statistical properties, in particular concerning the degree distribution and the clustering structure. These results suggests that models of online activity cannot discard the information carried by this multilayer representation of the system, and should account for the different processes generated by the different kinds of interactions. Secondly, our analysis unveils the presence of statistical regularities among the different events, suggesting that the non-trivial topological patterns that we observe may represent universal features of the social dynamics on online social networks during exceptional events.

Keywords: multilayer, social networks, complex networks, exceptional events, big data

1. Introduction

The advent of online social platforms and their usage in the last decade, with exponential increasing trend, made possible the analysis of human behavior with an unprecedented volume of data. To a certain extent, online interactions represent a good proxy for social interactions and, as a consequence, the possibility to track the activity of individuals in online social networks allows one to investigate human social dynamics [1].

More specifically, in the last years an increasing number of researchers focused on individual's activity in Twitter, a popular microblogging social platform with about 302 millions active users posting, daily, more than 500 millions messages (i.e., *tweets*) in 33 languages¹. In traditional social science research the size of the population under investigation is very small, with increasing costs in terms of human resources and funding. Conversely, monitoring Twitter activity, as well as other online social platforms as Facebook and Foursquare to cite just some of them, dramatically reduces such costs and allows to study a larger population sample, ranging from hundreds to millions of

¹<https://about.twitter.com/company>.

individuals [2], within the emerging framework of computational social science [3].

The analysis of Twitter revealed that online social networks exhibit many features typical of social systems, with strongly clustered individuals within a scale-free topology [4]. Twitter data [5] has been used to validate Dunbar's theory about the theoretical cognitive limit on the number of stable social relationships [6, 7]. It has been shown that individuals tend to share ties within the same metropolitan region and that non-local ties distance, borders and language differences affect their relationships [8]. Many studies were devoted to determine which and how information flows through the network [9–12], as well as to understand the mechanisms of information spreading—e.g., as in the case of viral content—to identify influential spreaders and comprehend their role [13–17]. Attention has also been given to investigate social dynamics during emergence of protests [18], with evidences of social influence and complex contagion providing an empirical test to the recruitment mechanisms theorized in formal models of collective action [19].

Twitter allows users to communicate through small messages, using three different actions, namely mentioning, replying and retweeting. While some evidences have shown that users tend to exploit in different ways the actions made available by the Twitter platform [20], such differences have not been quantified so far. In this work, we analyze the activities of users from a new perspective and focus our attention on how individuals interact during exceptional events.

In our framework, an exceptional event is a circumstance not likely in everyday news, limited to a short amount of time—typically ranging from hours to a few days—that causes an exceptional volume of tweets, allowing to perform a significant statistical analysis of social dynamics. It is worth mentioning that fluctuations in the number of tweets, mentions, retweets, and replies among users may vary from tens up to thousands in a few minutes, depending on the event. A typical example of exceptional event is provided by the discovery of the Higgs boson in July 2012 [21], one of the greatest events in modern physics.

We use empirical data collected during six exceptional events of different type, to shed light on individual dynamics in the online social network. We use social network analysis to quantify the differences between mentioning, replying and retweeting in Twitter and, intriguingly, our findings reveal universal features of such activities during exceptional events.

2. Materials and Methods

2.1. Material

It has been recently shown that the choice of how to gather Twitter data may significantly affect the results. In fact, data obtained from a simple backward search tend to over-represents more central users, not offering an accurate picture of peripheral activity, with more relevant bias for the network of mentions [14]. Therefore, we used the streaming Application Programming Interface (API) made available by Twitter, to collect all messages posted on the social network satisfying a set of temporal and semantic constraints. More specifically, we made use of the public

streaming API² subjected to filters (keywords, hashtags or a combination of both). If the flow of tweets corresponding to the filter is smaller than 1% of the total flow on Twitter, then all tweets satisfying the filters are obtained, otherwise a warning reporting the number of missed tweets is received.

We consider different exceptional events because of their importance in different subjects, from politics to sport. More specifically, we focus on the Cannes Film Festival in 2013³ (Cannes2013), the discovery of the Higgs boson in 2012⁴ [21] (HiggsDiscovery2012), the 50th anniversary of Martin Luther King's famous public speech "I have a dream" in 2013⁵ (MLKing2013), the 14th IAAF World Championships in Athletics held in Moscow in 2013⁶ (MoscowAthletics2013), the "People's Climate March"—a large-scale activist event to advocate global action against climate change—held in New York in 2014⁷ (NYClimateMarch2014) and the official visit of US President Barack Obama in Israel in 2013⁸ (ObamaInIsrael2013).

For each event, we collected tweets sent between a starting time t_i and a final time t_f containing at least one keyword or hashtag, as specified in **Table 1**. For almost all events, we have chosen keywords and hashtag that are very specific, reducing the amount of noise (i.e., tweets that are not related to the event although they satisfy our filters). In the case of the visit of Barack Obama in Israel in 2013 we have included the more generic keyword "peace," because in this specific context it was relevant for gathering data. However, it is worth anticipating here that our results show that the (unknown) amount of noise in this dataset did not alter the salient statistical features of the dataset.

Finally, we report that in a few cases we complemented a dataset by including tweets obtained from the search API (at most 5% of tweets with respect to the whole dataset) and that in the worst cases, the flow of streaming API was limited causing a loss of less than 0.5% of tweets.

2.2. Methods

To understand the dynamics of Twitter user interactions during these exceptional events, we reconstruct, for each event, a network connecting users on the basis of the retweets, mentions and replies they have been the subject or object of. In the literature on Twitter data what is usually built is the network based on the follower-followee relationships between users [4, 8, 9]. However, this kind of network only captures users' declared relations and it does not provide a good proxy for the actual interactions between them. Users, in fact, usually follow hundreds of accounts whose tweets appear in their news feed, even if there is no real interaction with the majority of those individuals. Therefore, to capture the social structure emerging from these interactions we build instead a network based on the exchanges between users, which can be deduced from the tweets that they

²<https://dev.twitter.com/streaming/public>.

³https://en.wikipedia.org/wiki/2013_Cannes_Film_Festival.

⁴https://en.wikipedia.org/wiki/Higgs_boson#Discovery_of_candidate_boson_at_CERN.

⁵https://en.wikipedia.org/wiki/I_Have_a_Dream.

⁶https://en.wikipedia.org/wiki/2013_World_Championships_in_Athletics.

⁷https://en.wikipedia.org/wiki/People's_Climate_March.

⁸https://en.wikipedia.org/wiki/List_of_presidential_trips_made_by_Barack_Obama#2013.

TABLE 1 | Information about events used in this work.

Dataset	Starting date	Ending date	Keywords
Cannes2013	06 May 2013	03 Jun 2013	cannes film festival,cannes, canneslive
	05:23:49 GMT	03:48:26 GMT	#cannes2013, #festivalcannes, #palmdor
HiggsDiscovery2012	30 Jun 2012	10 Jul 2012	lhc, cern, boson, higgs
	21:11:19 GMT	20:59:56 GMT	
MLKing2013	25 Aug 2013	02 Sep 2013	Martin Luther King
	13:41:36 GMT	08:16:21 GMT	#ihaveadream
MoscowAthletics2013	05 Aug 2013	19 Aug 2013	mos2013com, moscow2013, mosca2013
	09:25:46 GMT	12:35:21 GMT	moscu2013, #athletics
NYClimateMarch2014	18 Sep 2014	22 Sep 2014	peopleclimatemarch, peoplesclimate
	22:46:19 GMT	04:56:25 GMT	marciaxilclima, climate2014
ObamaInIsrael2013	19 Mar 2013	03 Apr 2013	obama, israel
	15:56:29 GMT	21:24:34 GMT	palestina, peace

Note that starting and ending dates reported here consider only tweets where users perform a social action, i.e., tweets without mentions, replies or retweets are not considered.

produce. In particular, there are three kinds of interactions that can take place on Twitter and that we will focus on:

- A user can **retweet** (RT) another user’s tweet. This means that the user is endorsing a piece of information shared by the other user, and is rebroadcasting it to her/his own followers.
- A user can **reply** (RP) to another user’s tweet. This represents an exchange from a user to another as a reaction of the information contained in a user’s tweet.
- A user can **mention** (MT) another user in a tweet. This represents an explicit share of a piece of information with the mentioned user.

A fourth kind of possible interaction is to favourite a user’s tweet, which represents a simple endorsement of the information contained in the tweet, without rebroadcasting. However, we do not have this kind of information for this dataset and therefore we do not consider this kind of interaction.

As just discussed, each kind of activity on Twitter (retweet, reply, and mention) represents a particular kind of interaction between two users. Therefore, an appropriate framework to capture the overall structure of these interactions without loss of information about the different types is the framework of multilayer networks [22–27]. More specifically, in the case under investigation the more appropriate model is given by edge-colored graphs, particular multilayer networks

TABLE 2 | Number of nodes and edges of the network corresponding to each event considered in this study.

Event	Aggregate	RT	RP	MT
Cannes2013	N = 514,328	337,089	85,414	91,825
	E = 700,492	490,268	82,952	127,272
HiggsDiscovery2012	N = 747,659	434,687	167,385	145,587
	E = 817,877	542,808	122,761	152,308
MLKing2013	N = 346,069	286,227	24,664	35,178
	E = 339,143	288,543	18,157	32,443
MoscowAthletics2013	N = 103,319	73,377	11,983	17,959
	E = 144,591	102,842	12,768	28,981
NYClimateMarch2014	N = 115,284	94,300	7,900	13,084
	E = 239,935	213,158	8,038	18,739
ObamaInIsrael2013	N = 2,641,052	1,443,929	737,353	459,770
	E = 2,926,777	1,807,160	586,074	533,543

The second column reports the total number of nodes and edges, corresponding to a network in which information is aggregated. The last three columns report the number of active nodes and edges per layer. A node is considered active on a given layer if the corresponding user is the subject or the object of the corresponding kind of interaction.

where a color is assigned to different relationships—i.e., the edges—among individuals defining as many layers as the number of colors. We refer to Kivelä et al. [28] and Boccaletti et al. [29] for thorough reviews about multilayer networks.

Here, for each event, we build a multilayer network composed by $L = 3$ layers {RT,RP,MT}, corresponding to the three actions that users can perform in Twitter, and N nodes, being N the number of Twitter users interacting in the context of the given event. A directed edge between user i and user j on the RT layer is assigned if i retweeted j . Similarly, an edge exists on RP layer if user i replied to user j , and on MT layer if i mentioned j . An illustrative example is shown in **Figure 1**.

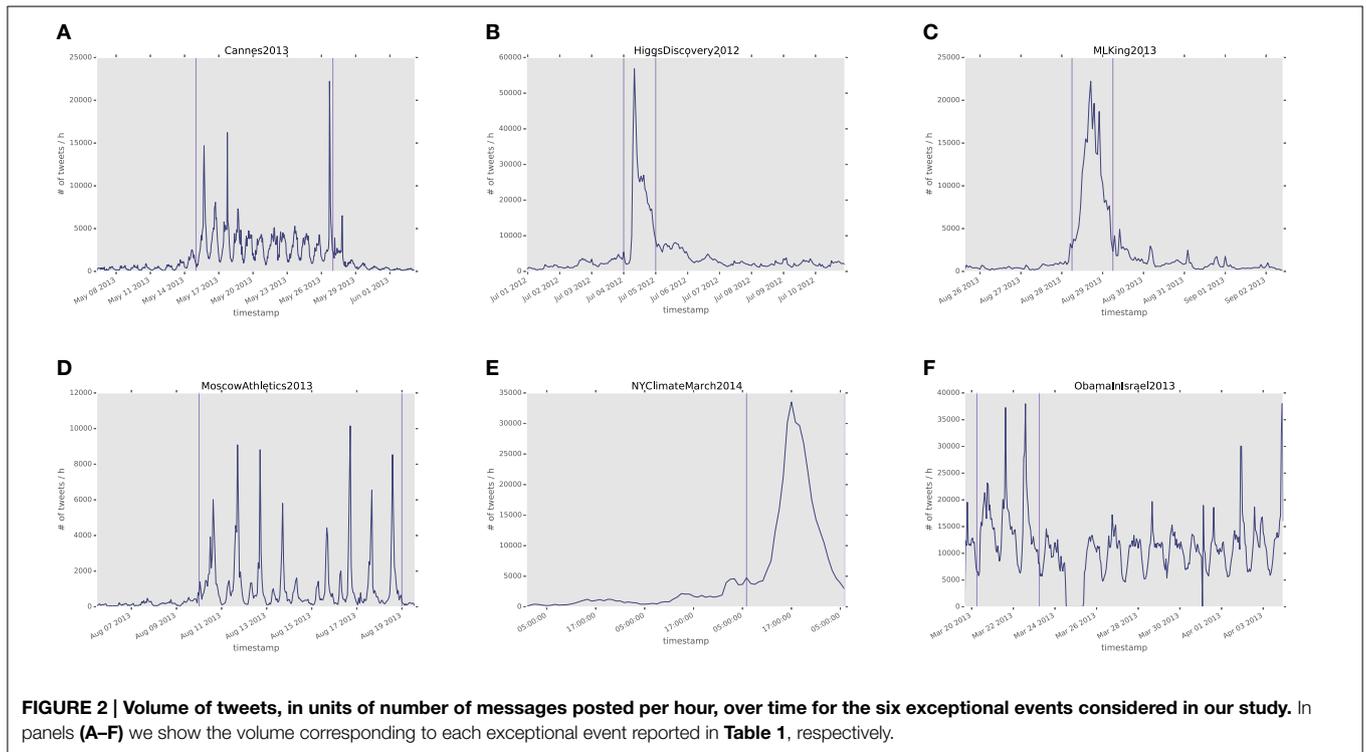
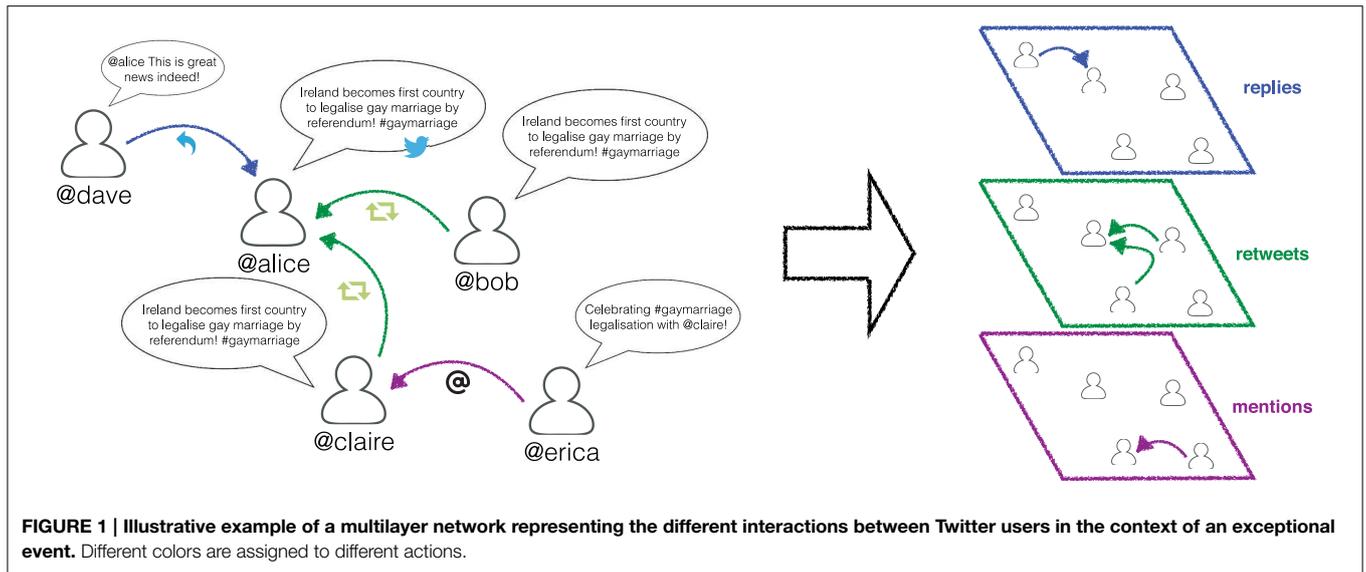
Details about the number of nodes and edges characterizing each event are reported in **Table 2**. We can observe that the number of nodes and edges can vary importantly across events and across layers, but for each event and each interaction type the size of the corresponding networks is sufficient to allow a statistically significant analysis of the data.

3. Results

In the following we present an analysis of the networks introduced in the previous section, which is oriented at exploring two different but complementary questions.

Firstly we want to know if, within one same event, the three kinds of interactions produce different network topologies. To this aim, we consider basic multilayer and single-layer network descriptors relevant to characterize social relationships, and we study how they vary when considering different layers.

Secondly, we want to unveil if different exceptional events present any common pattern regarding users interactions. As shown in **Figure 2**, the temporal pattern of the different events



considered in our study presents highly heterogeneous profiles. Some events are, in fact, limited to one day or only to a few hours, whereas others span over a week or more, and the profile of tweets volume varies accordingly. However, despite of these differences, do the user interactions that take place during these events present any common feature?

3.1. Edge Overlap Across Layers

To understand if the kinds of interaction produce similar networks or not, we analyze if users interact similarly with

each other regardless of the type of activity (retweet, reply, or mention), or not. This information can be obtained by calculating the edge overlap [26, 30] between each pair of layers. However, when the number of edges is very heterogeneous across layers, a more suitable descriptor of edge overlap is given by:

$$o_{\alpha\beta} = \frac{|E_{\alpha} \cap E_{\beta}|}{\min(|E_{\alpha}|, |E_{\beta}|)}, \quad (1)$$

where E_{α} (E_{β}) is the set of edges belonging to layer α (β) and $|\cdot|$ indicates the cardinality of the set. This measure quantifies

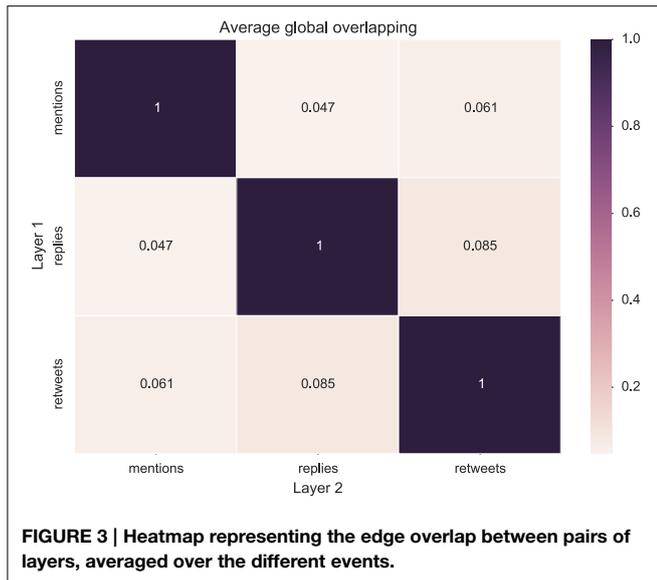


FIGURE 3 | Heatmap representing the edge overlap between pairs of layers, averaged over the different events.

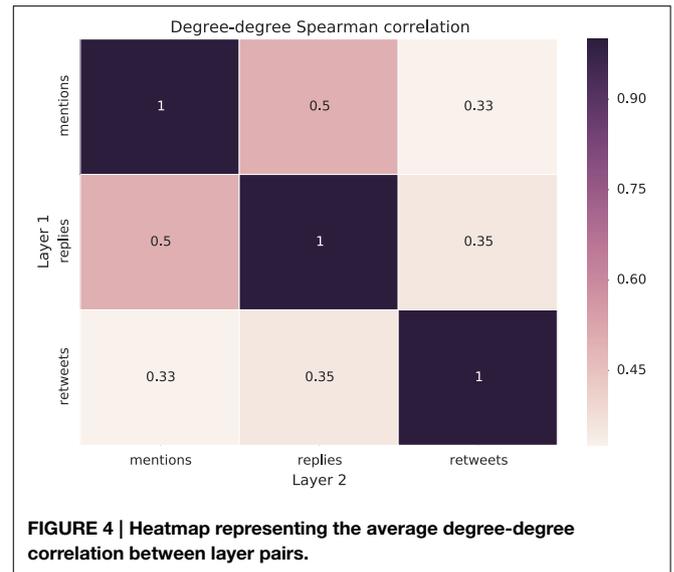


FIGURE 4 | Heatmap representing the average degree-degree correlation between layer pairs.

TABLE 3 | Number of nodes and edges of the network corresponding to each event considered in this study.

Layer pair	Edge overlap	Degree-degree correlation
MT-RP	0.05 ± 0.04	0.50 ± 0.12
MT-RT	0.06 ± 0.03	0.33 ± 0.08
RP-RT	0.08 ± 0.04	0.35 ± 0.10

the proportion of pair-wise interactions—represented by the edges—that are common to two different layers. Because, as shown in **Table 2**, the number of edges can vary largely on the different layers, the normalization is given by the cardinality of the smallest set of edges, to avoid biases resulting from the size difference. The results are reported in **Figure 3**. Each value is obtained by averaging over the different events. The standard deviations are not shown in the figure for the sake of clarity, but are reported in **Table 3**. We see that, for every couple of layers, (α, β) , $\rho_{\alpha\beta} \ll 1$. This result indicates that different layers contain different pairwise interactions, i.e., the users that we retweet are not necessarily the same that we mention or we reply to, for example. This result suggests that considering the different activities separately might be very relevant in order to understand human interaction dynamics on Twitter.

3.2. Degree-degree Correlations Across Layers

In this section, we study the degree connectivity of users, the most widely studied descriptor of the structure of a network. We focus in particular on the in-degree $k_{i,\alpha}$, which quantifies the number of users who interacted with user i on layer α ($\alpha = RT, RP, \text{ and } MT$). This is the simplest measure of the importance of the user in the network.

First, we explore if users have the same connectivity on the different layers, or not, i.e., if the users consistently have the same degree of importance on all the layers, or not. To this aim, we compute the Spearman’s rank correlation coefficient [31] between the in-degree of users on one layer and their in-degree on a different layer, for each pair of layers. The results,

averaged across the different events, are reported in **Figure 4**, with statistical details reported in **Table 3**. The value of two degree-degree correlations out of three is about 0.35, and the third—and highest—correlation is 0.5. This means that users tend to have different in-degree values on the different layers, i.e., a highly retweeted user is most likely not to be mentioned or replied to by as many users. This result suggests that the different types of interaction might produce different networks and should be considered separately in realistic modeling of individual dynamics.

3.3. Degree Distribution Per Layer

Building on the result discussed in the previous section, we also explore, for each event, the distribution of the in-degree on the different layers, separately. Intriguingly, for each layer, we find that the empirical distributions corresponding to all the exceptional events present very similar shape, as shown in **Figure 5**. This result suggests that individuals’ communications on Twitter present some universal characteristics across very different types of events.

The in-degree, shown in **Figure 5**, exhibits a power-law distribution for about three order of magnitudes. To validate our observation, we fit a power law to each distribution following a methodology similar to the one introduced in Clauset et al. [32]. By noticing that the in-degree is a discrete variable, we estimate the scaling exponent of a discrete power law for each empirical distribution. The goodness of fit is estimated by using the Chi Square test [33]. We find that the null hypothesis that the data is described by a discrete power law is accepted for all empirical distributions with a confidence level of 99%. We have tested other hypotheses, by considering other distributions with fat tails such as lognormal, exponential, Gumbel’s extreme values, and Poisson. In the cases where the null hypothesis is accepted with the same confidence level, we used the Akaike information criterion (AIC) [34, 35] to select the best model. It is worth remarking that, in all cases, we find that the power law provides the best description of the data.

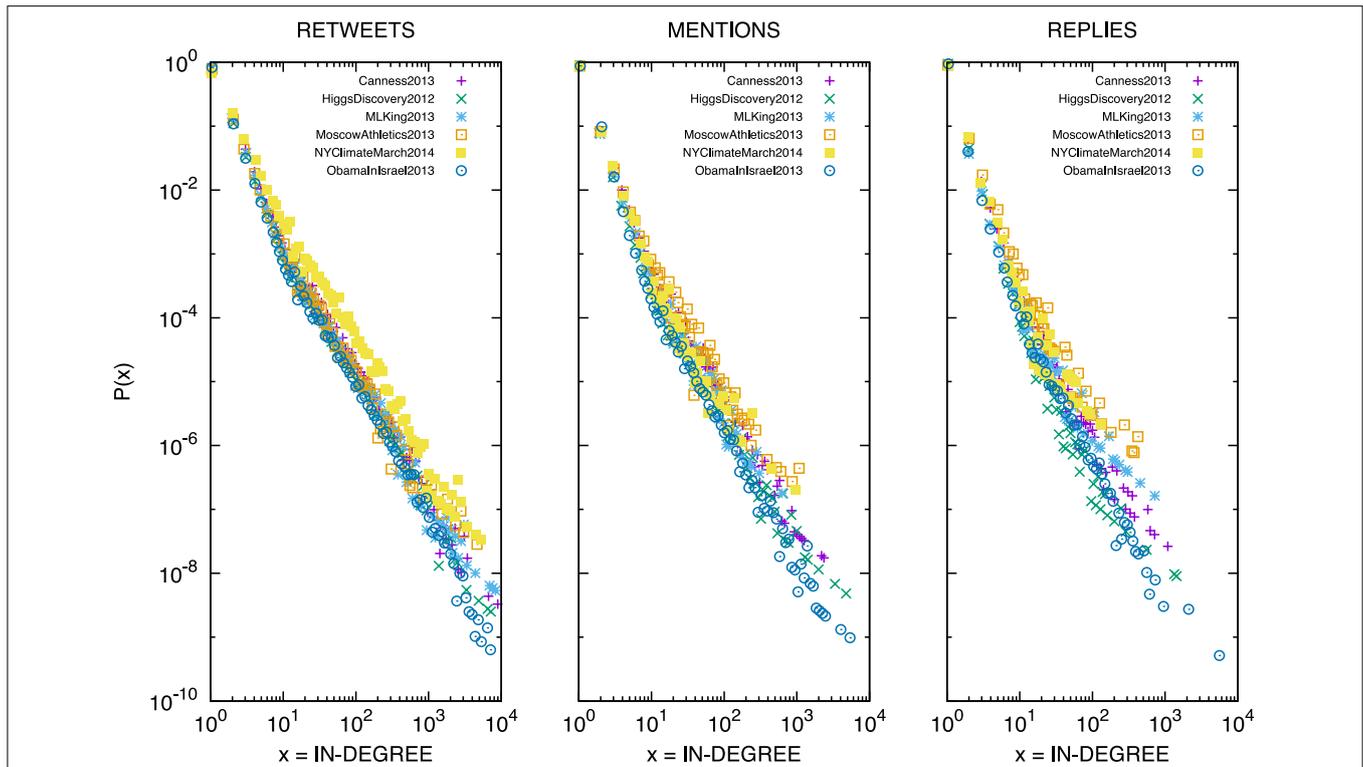


FIGURE 5 | Distribution of the in-degree for each event considered in this study (encoded by points with different shape and color) and each layer: retweets (left), mentions (center), and replies (right).

Power-law distributions of the degree have been found in a large variety of empirical social networks [36]. Here, the main finding of our results is that each kind of interaction presents a different scaling exponent. To show this, in **Figure 6** we report three notched box plots, each corresponding to a different layer and including the information about the different events. Notched box plots present a contraction around the median, whose height is statistically important: if the notches of two boxes do not overlap, this offers evidence of a statistically significant difference between the two medians. This is indeed the case in **Figure 6**, meaning that the median scaling exponent of the in-degree distribution of each of the three layer is different from the exponent characterizing the in-degree distribution of the other layers. The fact that the in-degree distributions corresponding to the different types of interaction are characterized by different scaling exponents indicates that the dynamics of each type of interaction in Twitter should be modeled as a distinct process, and that existing models of Twitter activity that do not take into account this fact should be carefully rethought.

3.4. Average Clustering Per Layer

Lastly, for each layer separately, we calculate the average clustering coefficient of the corresponding network. This is a measure of the transitivity of the observed interactions, and constitutes an important metric to characterize social networks

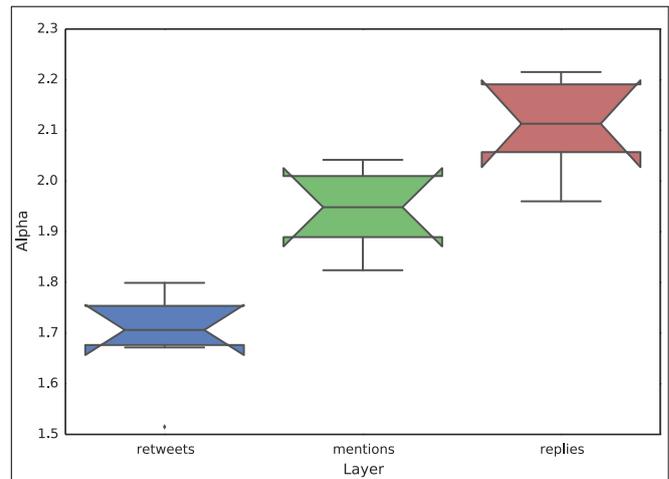
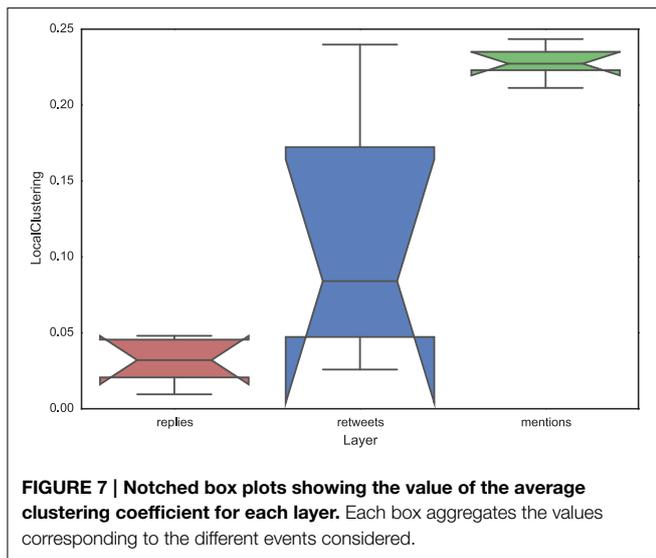


FIGURE 6 | Notched box plots showing the value of the scaling exponent of the in-degree distribution for each layer. Each box aggregates the values corresponding to the different events considered. Notched box plots present a contraction around the median, whose height is statistically important: if the notches of two boxes do not overlap, this offers evidence of a statistically significant difference between the two medians. This is the case here, meaning that the median scaling exponent of the in-degree distribution of each of the three layer is different from the exponent characterizing the in-degree distribution of the other layers.



[37]. In particular, for each event and each layer, we compute the average local clustering coefficient defined by:

$$\bar{C} = \frac{1}{N} \sum_{i=1}^N C_i, \tag{2}$$

where

$$C_i = \frac{2|\{e_{jk} : v_j, v_k \in N_i, e_{jk} \in E\}|}{k_i(k_i - 1)}, \tag{3}$$

where e_{jk} indicates the edge between users j and k . We show in **Figure 7** the values of the clustering coefficient using three notched box plots, each corresponding to a different layer and including the information about the different events. The mention network has the highest clustering level, whereas the reply network has the lowest one. The clustering level of the retweet network is the most variable across events, however the three medians are again different because the notches do not overlap. This result is a further confirmation that the three layers, and therefore the three types of interaction that they represent, form different network topologies

References

1. Centola D. The spread of behavior in an online social network experiment. *Science* (2010) **329**:1194–7. doi: 10.1126/science.1185231
2. Borge-Holthoefer J, Baños RA, González-Bailón S, Moreno Y. Cascading behaviour in complex socio-technical networks. *J Complex Netw.* (2013) **1**:3–24. doi: 10.1093/comnet/cnt006
3. Lazer D, Pentland AS, Adamic L, Aral S, Barabasi AL, Brewer D, et al. Life in the network: the coming age of computational social science. *Science* (2009) **323**:721. doi: 10.1126/science.1167742
4. Kwak H, Lee C, Park H, Moon S. What is twitter, a social network or a news media? In: *Proceedings of the 19th International Conference on World Wide Web*. Raleigh, NC: ACM (2010). pp. 591–600.

and that the dynamical processes producing them are thus distinct.

4. Discussion

In this paper we analyze six datasets consisting of Twitter conversations surrounding distinct exceptional events. The considered events span over very different topics: entertainment, science, commemorations, sports, activism, and politics. Our results show that, despite the different fluctuations in time and in volume, there are some statistical regularities across the different events. In particular, we find that the in-degree distribution of users and the clustering coefficient in each of the three layers (representing interactions based on retweet, replies, and mentions, respectively) are the same across the six different events. Our first conclusion is therefore that users behavior on Twitter—during exceptional events—presents some universal patterns.

Secondly, we show that different types of interactions between users on Twitter (retweeting, replying, and mentioning) generate networks presenting different topological characteristics. These differences were captured making use of the multilayer network framework: instead of discarding the information contained in the tweets regarding how users interact, we use this information to build a more complete representation of the system by means of three layers, each representing a different type of interaction. The fact that networks corresponding to different layer present different statistical properties is an important hint for models aiming at reproducing human behavior in online social networks. Our results indicate that, to faithfully represent how users interact, these models cannot be based on an aggregated view of the network and should account for all the different processes taking place in the system, separately.

Acknowledgments

AA and MD were supported by the European Commission FET-Proactive project PLEXMATH (Grant No. 317614) and the Generalitat de Catalunya 2009-SGR-838. AA also acknowledges financial support from the ICREA Academia, James S. McDonnell Foundation and MINECO FIS2012-38266. EO is supported by James S. McDonnell Foundation.

5. Gonçalves B, Perra N, Vespignani A. Modeling users' activity on twitter networks: validation of dunbar's number. *PLoS ONE* (2011) **6**:e22656. doi: 10.1371/journal.pone.0022656
6. Dunbar RI. Neocortex size as a constraint on group size in primates. *J Hum Evol.* (1992) **22**:469–93. doi: 10.1016/0047-2484(92)90081-J
7. Dunbar RI. Foundations in social neuroscience. In: Cacioppo JT, Berntson GG, Adolphs R, Carter CS, Davidson RJ, McClintock M, et al. editors. *The Social Brain Hypothesis*. Vol. 5. Cambridge, MA: The MIT Press (2002), p. 69.
8. Takhteyev Y, Gruzd A, Wellman B. Geography of twitter networks. *Soc Netw.* (2012) **34**:73–81. doi: 10.1016/j.socnet.2011.05.006
9. Java A, Song X, Finin T, Tseng B. Why we twitter: understanding microblogging usage and communities. In: *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*. San Jose, CA: ACM (2007). pp. 56–65.

10. Yang J, Counts S. Predicting the speed, scale, and range of information diffusion in twitter. In: *Proceedings of the Fourth International Conference on Weblogs and Social Media*. Vol. 10. Washington, DC (2010), pp. 355–8.
11. Wu S, Hofman JM, Mason WA, Watts DJ. Who says what to whom on twitter. In: *Proceedings of the 20th International Conference on World Wide Web*. Hyderabad: ACM (2011). pp. 705–14.
12. Myers SA, Zhu C, Leskovec J. Information diffusion and external influence in networks. In: *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Beijing: ACM (2012). pp. 33–41.
13. Bakshy E, Hofman JM, Mason WA, Watts DJ. Everyone's an influencer: quantifying influence on twitter. In: *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*. Kowloon: ACM (2011). pp. 65–74.
14. González-Bailón S, Wang N, Rivero A, Borge-Holthoefer J, Moreno Y. Assessing the bias in samples of large online networks. *Soc Net.* (2014) **38**:16–27. doi: 10.1007/s10955-012-0595-6
15. Borge-Holthoefer J, Meloni S, Gonçalves B, Moreno Y. Emergence of influential spreaders in modified rumor models. *J Stat Phys.* (2013) **151**:383–93. doi: 10.1007/s10955-012-0595-6
16. González-Bailón S, Borge-Holthoefer J, Moreno Y. Broadcasters and hidden influentials in online protest diffusion. *Am Behav Sci.* (2013) **57**:943–65. doi: 10.1177/0002764213479371
17. Baños RA, Borge-Holthoefer J, Moreno Y. The role of hidden influentials in the diffusion of online information cascades. *EPJ Data Sci.* (2013) **2**:1–16. doi: 10.1140/epjds18
18. Borge-Holthoefer J, Rivero A, García I, Cauhé E, Ferrer A, Ferrer D, et al. Structural and dynamical patterns on online social networks: the spanish may 15th movement as a case study. *PLoS ONE* (2011) **6**:e23883. doi: 10.1371/journal.pone.0023883
19. González-Bailón S, Borge-Holthoefer J, Rivero A, Moreno Y. The dynamics of protest recruitment through an online network. *Sci Rep.* (2011) **1**:197. doi: 10.1038/srep00197
20. Conover M, Ratkiewicz J, Francisco M, Gonçalves B, Menczer F, Flammini A. Political polarization on twitter. In: *Fifth International AAAI Conference on Weblogs and Social Media*. Barcelona (2011).
21. De Domenico M, Lima A, Mougél P, Musolesi M. The anatomy of a scientific rumor. *Sci Rep.* (2013) **3**:2980. doi: 10.1038/srep02980
22. Cardillo A, Gómez-Gardeñes J, Zanin M, Romance M, Papo D, del Pozo F, et al. Emergence of network features from multiplexity. *Sci Rep.* (2012) **3**:1344. doi: 10.1038/srep01344
23. De Domenico M, Solé-Ribalta A, Cozzo E, Kivelä M, Moreno Y, Porter MA, et al. Mathematical formulation of multilayer networks. *Phys Rev X.* (2013) **3**:041022. doi: 10.1103/PhysRevX.3.041022
24. Nicosia V, Bianconi G, Latora V, Barthelemy M. Growing multiplex networks. *Phys Rev Lett.* (2013) **111**:058701. doi: 10.1103/PhysRevLett.111.058701
25. Gallotti R, Barthelemy M. Anatomy and efficiency of urban multimodal mobility. *Sci Rep.* (2013) **4**:6911. doi: 10.1038/srep06911
26. De Domenico M, Nicosia V, Arenas A, Latora V. Structural reducibility of multilayer networks. *Nat Commun.* (2015) **6**:6864. doi: 10.1038/ncomms7864
27. De Domenico M, Solé-Ribalta A, Omodei E, Gómez S, Arenas A. Ranking in interconnected multilayer networks reveals versatile nodes. *Nat Commun.* (2015) **6**:6868. doi: 10.1038/ncomms7868
28. Kivelä M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA. Multilayer networks. *J Complex Netw.* (2014) **2**:203–71. doi: 10.1093/comnet/cnu016
29. Boccaletti S, Bianconi G, Criado R, Del Genio CI, Gómez-Gardeñes J, Romance M, et al. The structure and dynamics of multilayer networks. *Phys Rep.* (2014) **544**:1–122. doi: 10.1016/j.physrep.2014.07.001
30. Battiston F, Nicosia V, Latora V. Structural measures for multiplex networks. *Phys Rev E* (2014) **89**:032804. doi: 10.1103/PhysRevE.89.032804
31. Spearman C. The proof and measurement of association between two things. *Am J Psychol.* (1904) **15**:72–101. doi: 10.2307/1412159
32. Clauset A, Shalizi CR, Newman ME. Power-law distributions in empirical data. *SIAM Rev.* (2009) **51**:661–703. doi: 10.1137/070710111
33. Snedecor GW, Cochran WG. *Statistical Methods*. Iowa City, IA: Iowa State University Press (1967).
34. Akaike H. Information theory and an extension of the maximum likelihood principle. In: *2nd International Symposium on Information Theory*. Tsahkadsor: Akadémiai Kiadó (1973). pp. 267–81.
35. Akaike H. A new look at the statistical model identification. *Automat Control IEEE Trans.* (1974) **19**:716–23. doi: 10.1109/TAC.1974.1100705
36. Caldarelli G. *Scale-free Networks: Complex Webs in Nature and Technology*. Oxford, UK: OUP Catalogue (2007).
37. Watts DJ, Strogatz SH. Collective dynamics of small-world networks. *Nature* (1998) **393**:440–2. doi: 10.1038/30918

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Omodei, De Domenico and Arenas. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Digital daily cycles of individuals

Talayah Aledavood^{1*}, Sune Lehmann^{2,3} and Jari Saramäki¹

¹ Department of Computer Science, Aalto University School of Science, Espoo, Finland, ² DTU Compute, Technical University of Denmark, Lyngby, Denmark, ³ The Niels Bohr Institute, University of Copenhagen, Copenhagen, Denmark

Humans, like almost all animals, are phase-locked to the diurnal cycle. Most of us sleep at night and are active through the day. Because we have evolved to function with this cycle, the circadian rhythm is deeply ingrained and even detectable at the biochemical level. However, within the broader day-night pattern, there are individual differences: e.g., some of us are intrinsically morning-active, while others prefer evenings. In this article, we look at digital daily cycles: circadian patterns of activity viewed through the lens of auto-recorded data of communication and online activity. We begin at the aggregate level, discuss earlier results, and illustrate differences between population-level daily rhythms in different media. Then we move on to the individual level, and show that there is a strong individual-level variation beyond averages: individuals typically have their distinctive daily pattern that persists in time. We conclude by discussing the driving forces behind these signature daily patterns, from personal traits (morningness/eveningness) to variation in activity level and external constraints, and outline possibilities for future research.

OPEN ACCESS

Edited by:

Taha Yasseri,
University of Oxford, UK

Reviewed by:

Yougui Wang,
Beijing Normal University, China
Michael Szell,
Northeastern University, USA

*Correspondence:

Talayah Aledavood,
Department of Computer Science,
Aalto University, Otaniementie 17,
02150 Espoo, Finland
talayah.aledavood@aalto.fi

Specialty section:

This article was submitted to
Interdisciplinary Physics,
a section of the journal
Frontiers in Physics

Received: 17 July 2015

Accepted: 24 August 2015

Published: 07 October 2015

Citation:

Aledavood T, Lehmann S and
Saramäki J (2015) Digital daily cycles
of individuals. *Front. Phys.* 3:73.
doi: 10.3389/fphy.2015.00073

Keywords: circadian rhythms, electronic communication records, mobile phones, digital phenotyping, individual differences

1. Introduction

Almost all life on Earth is affected by the planet's 24-h period of rotation. Humans are no different; the rhythms of our lives are phase-locked with the diurnal cycle. Because our bodies have evolved to cope with the external environment, we have genetic circadian pacemaker circuits that intrinsically follow a period of approximately 24 h [the circadian period length may vary from one person to another, vary by age and there are known gender differences [1, 2]]. The operation of these circadian circuits manifests at various levels: biochemical, physiological, psychological, and in various markers from hormone levels to body temperature [3–6]. While our daily rhythms can be modulated by exogenous factors [e.g., decoupling alertness from the sleep/wake cycle [7]], there is a very strong endogenous component in these rhythms, as indicated by the persistence of a near-24 h rhythm in the absence of environmental cues or despite imposition of a non-24 h schedule [8, 9].

Within this broader pattern, however, there are substantial inter-individual differences. Such differences are apparent in the existence of *chronotypes*—morning types and evening types, those who go to bed early and those who find it difficult to wake up early. The traits of *morningness* and *eveningness* correlate with distinctive temporal patterns of physiological and psychological variables, such as body temperature and efficiency. They also appear to be linked to gender as well as personality traits; in particular, studies have shown weak negative correlations of morningness with extraversion and sociability [10, 11].

The daily rhythms that humans follow are visible in the digital records that are left in the wake of human online activity. Population-level and system-level daily rhythms can be observed in time variation of activity in Youtube, Twitter and Slashdot, and in frequency of edits in Wikipedia and OpenStreetMap [12–15]. They are also seen in the frequency of mobile telephone

calls [16, 17], and in traces of human mobility derived from mobile phone data [18–20]. But what do the circadian patterns displayed by activity levels in an online system actually reveal about human behavior? The behavior of an online system is determined by a number of factors: the day/night cycle, the function and purpose of the system in question (e.g., work-related emails mostly being sent during office hours, see below), the variation of behaviors of user groups (e.g., Wikipedia edits from multiple time zones), and, importantly, variation at the individual level.

In this paper, we discuss findings regarding the daily patterns in electronic records of human communication, along with results of analyses that illustrate such patterns in four different datasets. We start at the aggregate level, studying system-level average patterns and discuss the origins of the findings. From the system level, we will move on to the level of individuals, and focus on the variation that remains hidden within system-level averages: individual differences reflected in persistent, distinct daily activity patterns. This part confirms that earlier findings of persistent individual differences in a mobile telephone dataset [21] are general, and that persistent, distinct daily patterns of individuals are common to different communication channels. These findings are important in two ways: one is that in order to better understand human behavior, more focus is required on individual-level behavior. Second, showing that behavior of each individual persists in time opens up several new questions to better understand the reasons behind this persistence and how and why this persistence can be perturbed. We conclude by discussing the implications of these findings, and address future research questions from large-scale analysis of sleep habits of individuals with big data to daily activity patterns as part of digital phenotypes.

2. Daily Patterns at the Aggregate Level

2.1. Previous Work

Let us begin by discussing observations of digital daily cycles in different systems at the aggregate level, computed from digital records of communication and online activity. In every instance where the temporal variation of the activity levels in such systems is monitored, the result is a periodic pattern of activity on several time scales [22]. The longest scale is that of a calendar year, where special periods such as holidays can typically be distinguished (see, e.g., 17). Then there is a weekly cycle, where weekends typically differ from weekdays, and where there can be differences between weekdays as well [12–14, 17, 23]. Finally, there is a daily pattern which may significantly differ between different systems.

We stress that any observed system-level pattern rises out of the superposition of a multitude of individual patterns, and attributing system-level behavior to individuals would amount to an ecological fallacy. Therefore, interpreting what the system-level patterns represent remains a non-trivial task. Solving the problem of disentangling the superposition of daily patterns, however, may provide important information of the user population. A good example of this is Yasseri et al. [14], where the authors studied Wikipedia in various languages, and were able to infer the geographical spread of their editor base

from the assumption that the observed edit frequency cycles are a superposition of circadian patterns on different time zones. The method is based on the argument that Wikipedias in different languages exhibit universal daily patterns, with minima and maxima at around the same time of the day (when correcting for time zones).

Temporal patterns of activity have been studied for different online platforms. For example, in Yasseri et al. [15], the authors look at differences between editing patterns on OpenStreetMap, which is a geo-wiki, for two different cities (London and Rome). Circadian patterns of edits for the two cities have been compared to each other and to that of Wikipedia edits. The authors also followed changes in the circadian rhythms for each of the two cities over several years. In ten Thij et al. [24], daily and weekly patterns of Twitter activity in different languages have been studied and it has been shown that circadian patterns emerge for tweets in all the studied languages. In Noulas et al. [25], the authors have looked at data from Foursquare and found geotemporal rhythms in activity both for weekdays and weekends.

Analysis of aggregate-level daily cycles with geospatial information has been used in the context of cities and transport. As an example, in Toole et al. [26], the authors infer dynamic land use of different parts of a city based on temporal patterns of mobile phone activity in different locations. In Ahas et al. [19], temporal data is combined with location data from mobile phones. Comparing daily rhythms for different days of the week, the authors show a significant difference in mobility of suburban commuters in city of Tallinn on weekends as compared to work days. In Louail et al. [20], the authors investigate the daily rhythms of different Spanish cities in terms of spatiotemporal patterns of mobile phone usage, and show how the structure of hotspots, places of frequent usage, allow them to distinguish between different cities. Also, in Grauwin et al. [27], the authors study rhythms of mobile phone traffic records in three global cities in three different continents (London, New York, and Hong Kong). They look at daily patterns at the city level as well as at the local scale within each city and find similarities between cities in some features as well as distinctive patterns for each city for other features. In Dong et al. [28], Call Detail Record (CDR) data for a period of 5 months from Cote d'Ivoire is used to detect unusual crowd events and gatherings.

As a more applied and non-conventional example of the analysis of daily rhythms, in May 2014 a number of different news outlets (e.g., 29) described how an elaborate campaign run by Iranian hackers on social media, targeting American officials and figures, was revealed only after analysing the temporal patterns of three years of activity. The daily and weekly activity patterns of the hackers matched precisely the activity profile of Tehran (i.e., low activity at lunch hours of Tehran local time, and little or no activity on Thursdays and Fridays which are weekend days in Iran).

Finally, let us mention that electronic records contain evidence of daily/weekly patterns that go beyond activity rates. Using network analysis [17] show that when mobile telephone calls between individuals are aggregated to form networks, the structural features of those networks differ depending on the starting time of the aggregation process. In particular, weekends

differ from weekdays. It is probable that the explanation is that during weekends, communication is mainly targeted to close friends and relatives who reside within the dense core of one's egocentric network. At a smaller scale, in Aledavood et al. [21], the authors show that closest friends are frequently called in the evenings.

2.2. Results

In this work, we study three different datasets, one with calls, one with calls and text messages, and one containing email records [30]. For calls, we use the Reality Mining dataset [31], and another mobile phone dataset containing data from a small town in a European country with a population of around 8000 people, a subset of the data used in e.g., [32]. For the latter, we also study text messages. For all sets, we use 8-week slices. A summary of different sets can be found in **Table 1**. Preprocessing of the data is discussed in Section 5.

As the first step, we look at aggregated hourly event frequencies for each of the four different sets (**Figure 1**). It is clear that while the sleep/wake cycle is apparent in each set, there are also noticeable differences. Calls in the European town show a double-peaked daily curve, whereas the Reality Mining

TABLE 1 | Overview of the datasets used in this study.

Dataset	Participants	Active users	Total events
Reality mining call	87	47	14,187
Town call	1204	277	45,844
Town text	708	64	13,014
Email	2430	431	206,723

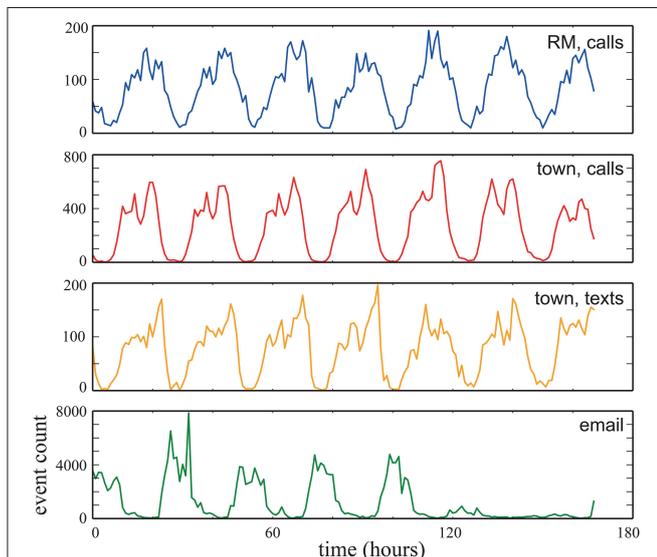


FIGURE 1 | Number of events per hour for each day of week in our datasets. This curve has been aggregated over the entire 8-week period. From top to bottom: calls in Reality Mining, calls and texts in small town, and emails. We observe strong diurnal patterns in all datasets; for the small town datasets there are also differences between calls and texts activity. The email dataset shows decreased activity during weekends.

data displays no such pattern. It is possible that this is due to different conventions; students in Boston can be expected to behave differently than people in a small European town. Note that for the Reality Mining data, time zone information is not available, so we have manually shifted them such that the lowest points correspond to night and there is a possibility that this estimate is inaccurate. However, this only affects the phase of the pattern, not its shape.

Interestingly, in both call datasets, the highest peak occurs on the fifth day (Friday). Also note the very low email activity level during the weekend in the email data. For email, time stamps are relative to some unknown t_0 , so the daily cycles appear shifted compared to the other datasets.

In **Figure 2** we focus on the difference between daily cycles the various datasets. Here, we plot the average daily patterns in each system on the third day of the week. Since there is no exact time zone information for Reality Mining and email datasets, we identified the third day of the week by assuming that two low-activity days correspond to the weekend. We also aligned the timelines by assuming that the lowest activity of the day occurs at 4 AM for all datasets. We then average over the third-day patterns across all 8 weeks in each set. As in Aledavood et al. [33], we find differences between the communication channels: for the small town dataset, the peak of text messages is later than that of calls. This is perhaps due to different nature of these channels; while getting calls in the late hours might not be appreciated, receiving text messages which are much less obtrusive is still acceptable.

3. Daily Patterns at the Level of Individuals

3.1. Previous Work

In Aledavood et al. [21], two present authors investigated individual-level daily cycles in mobile phone call data from 24 individuals (12 male and 12 female) over 18 months. The data collection was performed in a setting where the participants completed high school some months after the collection began,

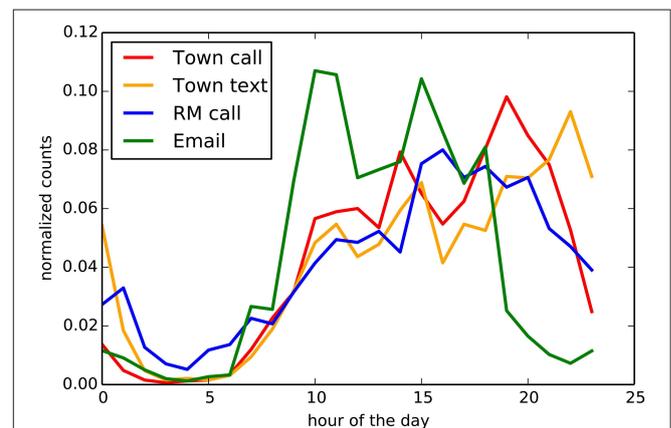
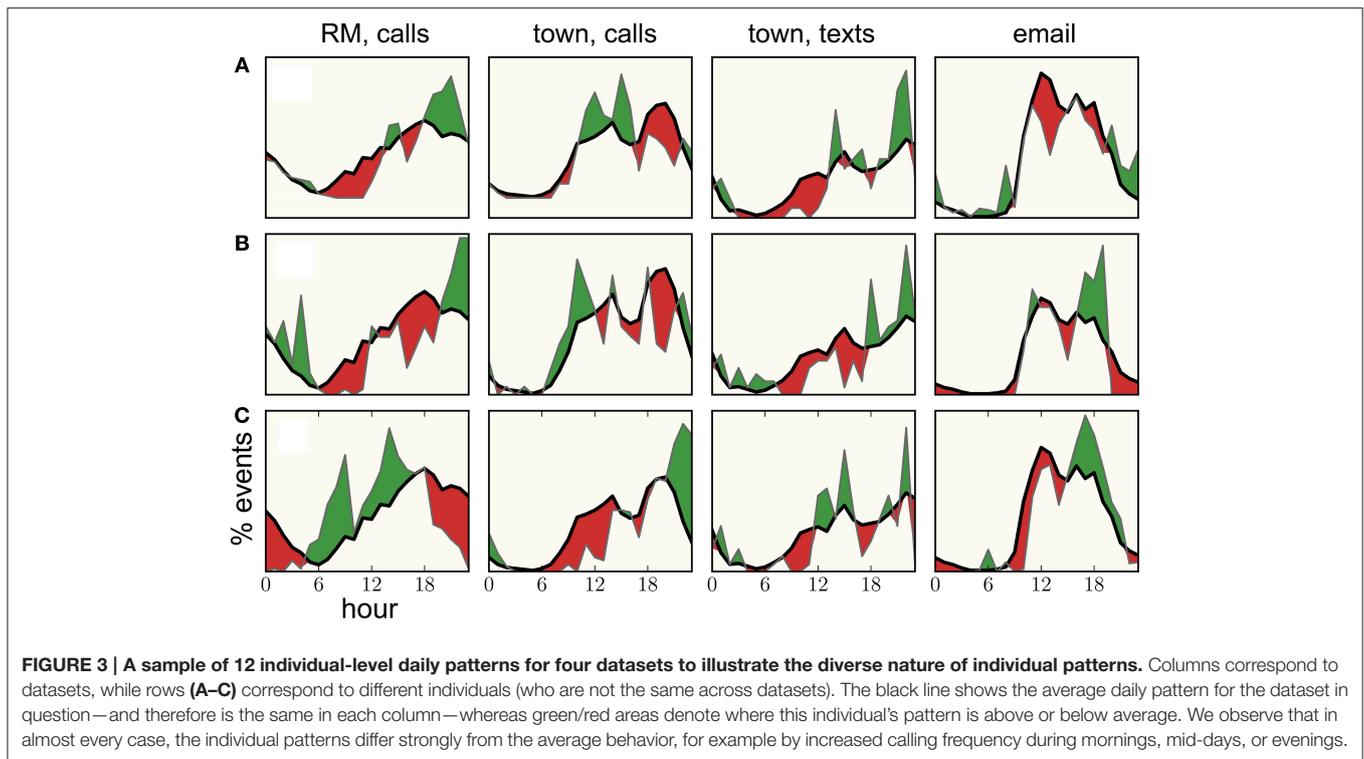


FIGURE 2 | The daily pattern in each of the datasets, computed as an average over all Wednesdays in the data. Colors are the same as in **Figure 1**. We observe distinct patterns across the various data channels. Email activity is early in the day, whereas (unobtrusive) text messages peak late at night.



and then started their first year at university, often in another city, or went to work. This design guaranteed a high turnover in their social networks [34], and provided an opportunity to study a major change in their life circumstances. Looking at individual-level daily call patterns, however, it was clear that there were persistent individual differences; each individual has their distinctive daily cycle despite social network turnover and changes in circumstances. This observation speaks in favor of intrinsic factors (such as the aforementioned chronotypes) dominating individual-level variations in daily patterns (see Section 4).

3.2. Results

Continuing the analysis of the four datasets, we first calculate for each set the daily patterns for each individual (“ego”) by counting the total number of events associated with the ego at each hour of the day through the whole 8 weeks. The counts are then normalized to one for each ego to yield that person’s daily activity pattern. As a reference, we also compute the average pattern over all egos from the normalized patterns.

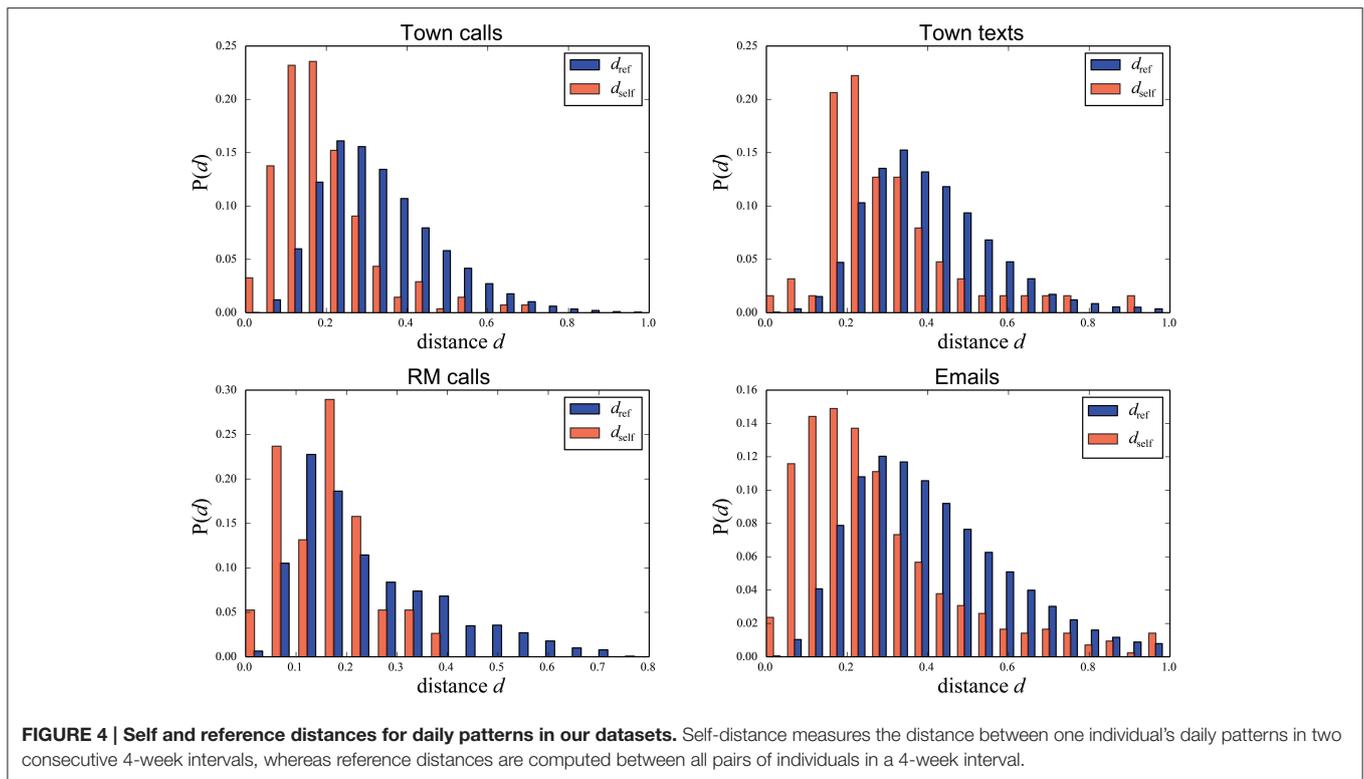
Figure 3 displays a sample of the individual-level daily patterns for each dataset. For each set, we have picked three egos to demonstrate individual differences; for each ego, their differences from the aggregated average are emphasized by red and green colors. For all datasets, we can observe clear variation between individuals. Considering the differences between the aggregate and individual daily cycles serves two purposes. While the average pattern in each dataset reveals general underlying mechanisms, the individual patterns show that each person has their own preferences for the timing of communication with

others. The daily communication cycles point at variation beyond morningness and eveningness: while individuals clearly have different sleep/wake cycles, they also have their specific patterns during their wakefulness periods.

Using the same methodology as Aledavood et al. [21] in order to study whether these daily patterns for each individual are persistent and thus characteristic for the individual, we divide the 8 weeks of data into two 4-week time intervals and use the Jensen–Shannon divergence to measure self and reference distances between patterns. A detailed explanation of these calculations can be found in the Section 5. The results are shown in **Figure 4**. We observe an effect similar to the findings in Aledavood et al. [21]: the daily patterns of individuals tend to be more similar to themselves in consecutive time intervals as compared to daily patterns of other individuals in the same time interval. This indicates that individuals have distinct daily patterns that retain their shapes in time. In other words, **Figure 4** shows that the individual differences seen in **Figure 3** are not just caused by random fluctuations: were fluctuations the reason for individual differences, each individual’s patterns in consecutive intervals would be equally similar or dissimilar to those of everyone else. As self-distances are on average lower, this is clearly not the case.

4. Discussion

Circadian rhythms have deep roots in human physiology, driven by the environment in which we live. These patterns manifest themselves in different ways at the individual and aggregate levels. There are diurnal patterns that are only visible at the



aggregate level in the overall frequencies of various phenomena that are rare or one-time events at the individual level: time of birth, heart attacks, suicides or committing unethical behavior [35–37]. To the contrary, the daily rhythms that we have focussed on here originate at the level of individuals, where they manifest as time-dependent event rates of e.g., digital communication.

What are the factors that determine an individual's daily rhythm as viewed through the lens of electronic records? The most obvious one is the sleep/wake cycle: we do not send emails or edit Wikipedia while asleep. This is known to be the central driver behind individual differences. First, individuals have different intrinsic chronotypes [morningness/eveningness tendencies [3]]. Second, the preferred duration of sleep also varies from one person to another [38]. Third, besides these intrinsic factors, external forcing such as different work schedules also have an effect on the sleep/wake cycle [39].

In addition to differences in the sleep/wake cycle, our alertness and propensity to sleep are distinct for each individual and vary throughout the day. Naturally, individuals go on average through fairly similar cycles of wakefulness and sleepiness, which may explain the qualitatively similar features of aggregate-level daily patterns across different systems. At the level of individuals, however, there are important differences, which are reflected in the observed daily patterns in digital records. As an example, a tired person might be less likely to write an important email or edit a Wikipedia article. Likewise, in addition to these intrinsic alertness cycles, one's daily schedule (work, commuting, etc.) plays a role by imposing constraints on the times when it is

possible to send emails or make calls. In terms of daily patterns of telephone calls, things are more complicated, because every call involves two individuals—a caller and a recipient. When calling, one must consider social norms and the availability of the other party.

Understanding which of the factors discussed above dominate the digital daily cycles of individuals and give rise to individual differences and persistent circadian patterns is a task that requires further attention. While the persistence of daily patterns appears to indicate that the intrinsic components (chronotypes, alertness cycles) do play a major role [21], external factors should also be of importance (see, e.g., 40). Further, it will be necessary to study whether individuals bound by (strong) social ties tend to synchronize their communication and availability.

While analysing digital records at the aggregate level can provide us invaluable population-level insights and help to replace or improve traditional survey or census methods [26, 41], studying the temporal fingerprints of individuals will unveil many new opportunities. As smartphones and other wearable devices are becoming ever more ubiquitous, they also increasingly provide high-velocity, high-volume data streams describing human behavior [42]. This data-collection capability makes these devices excellent tools for research, particularly within health, psychology and medicine, since smartphones allow researchers to study individual behavioral patterns [“digital phenotypes,” [43, 44] and their changes over time [45]. Monitoring an individual's digital behavioral patterns on different timescales is also an easy and inexpensive way for medical intervention, especially in the case of mental problems,

where there are fewer biomarkers than for other types of disease. Data from smartphones have already been used to monitor the time evolution of different measures that are known to be indicative of behavioral changes in patients, which makes daily monitoring and early intervention possible [46–48]. As an example, Faurholt-Jepsen et al. [49] suggest that data from mobile phones can be used as objective measure of symptoms of bipolar disorder.

Because the sleep/wake cycle is a dominant feature of circadian patterns, Big Data describing the digital daily cycles of large numbers of individuals might prove to be highly useful for sleep research. However, obtaining an accurate picture of the sleep times of individuals requires solving several non-trivial problems. While one does not send emails when asleep, emails are not necessarily a reliable proxy for awake-time; it is possible to be awake and not send emails. In this sense inferring the actual times of sleep from electronic records is challenging. This problem is made more severe by the ubiquitous burstiness in human dynamics [32, 50, 51]: broadly distributed inter-event times make the times from last observation to bed time (or from wake-up to first observation) highly unpredictable. Nevertheless, we believe that this is an important direction for future research.

Finally, a particularly promising source of data comes from large dedicated cell-phone based data collection efforts, focusing on collecting multiplex (face-to-face, telecommunication, online social networks) network data in a large, densely connected populations, e.g., [52]. Data from a single communication channel can be too sparse and noisy for obtaining accurate daily patterns; here, having a multiplex dataset can provide a great advantage since one can combine information from all data-channels to form a much more comprehensive picture of the activity of each person (e.g., for studying sleeping patterns). Furthermore, if the participants of the dataset are densely connected through social ties, it is also possible to investigate the significance of and correlations between the activity patterns of close personal relations using such a dataset. Finally, a dataset of this nature may function as a kind of “rosetta stone,” helping researchers determine the biases of each electronic dataset, and allowing us to understand to which extent telecommunication data or Twitter datasets with hundreds of millions of active users can be used to study the daily cycles of individuals.

References

- Schmidt C, Peigneux P, Cajochen C. Age-related changes in sleep and circadian rhythms: impact on cognitive performance and underlying neuroanatomical networks. *Front Neurol.* (2012) 3:118. doi: 10.3389/fneur.2012.00118
- Duffy JF, Cain SW, Chang AM, Phillips AJK, Münch MY, Gronfier C, et al. Sex difference in the near-24-hour intrinsic period of the human circadian timing system. *Proc Natl Acad Sci USA.* (2011) 108:15602–8. doi: 10.1073/pnas.1010666108
- Kerkhof GA. Inter-individual differences in the human circadian system: a review. *Biol Psychol.* (1985) 20:83–112. doi: 10.1016/0301-0511(85)90019-5
- Czeisler CA, Duffy JF, Shanahan TL, Brown EN, Mitchell JF, Rimmer DW, et al. Stability, precision, and near-24-hour period of the human circadian pacemaker. *Science* (1999) 284:2177–81. doi: 10.1126/science.284.5423.2177

5. Methods

5.1. Data Filtering

We have used 8-week time slices of all datasets. Filters have been applied to remove users who are inactive or whose activity is too low for producing meaningful information on daily patterns. In **Table 1**, the total number of participants means the total number of users who have at least one event during the study period of 8 weeks. For plotting aggregate-level patterns (**Figures 1, 2**), we have used data from all participants. The column “Active users” in the table represents the number of users who have at least one event per day on average (minimum 56 events in total); these have been used for calculating average daily patterns (**Figure 3**). For measuring persistence of daily patterns and calculating Jensen–Shannon divergence, we used a subset of active users who have at least one event in each of the two time intervals of 4 weeks.

5.2. Self and Reference Distances

In order to quantify the level of persistence of daily patterns for individuals, we compare the daily patterns of each ego for two consecutive 4-week time intervals. For this, we use the Jensen–Shannon divergence (JSD) and measure the distance of the daily patterns viewed as two probability distributions (P_1 and P_2). The JSD is calculated as follows: $JSD(P_1, P_2) = H(\frac{1}{2}P_1 + \frac{1}{2}P_2) - \frac{1}{2}[H(P_1) - H(P_2)]$, where $P_i = p(h)$ and $p(h)$ is the fraction of calls at each hour, $i = 1, 2$ indicates the time interval, and $H(P) = -\sum p(h) \log p(h)$ is the Shannon entropy. In order to compare these self-distances against a reference, we calculate a set of reference distances d_{ref} as the distances between the daily patterns of each ego and all other egos in the same time interval.

Author Contributions

TA, SL, and JS designed research. TA analyzed the data. TA, SL, and JS wrote the paper.

Acknowledgments

TA and JS acknowledge financial support from the Academy of Finland, project No. 260427. TA thanks Richard Darst for useful discussions.

- Panda S, Hogenesch J, Kay S. Circadian rhythms from flies to humans. *Nature* (2002) 417:329–35. doi: 10.1038/417329a
- Baehr EK, Revelle W, Eastman CI. Individual differences in the phase and amplitude of the human circadian temperature rhythm: with an emphasis on morningness-eveningness. *J Sleep Res.* (2000) 9:117–27. doi: 10.1046/j.1365-2869.2000.00196.x
- Folkard S, Hume KI, Minors DS, Waterhouse JM, Watson FL. Independence of the circadian rhythm in alertness from the sleep/wake cycle. *Nature* (1985) 313:678–9. doi: 10.1038/313678a0
- Kleitman N. *Sleep and Wakefulness*. Chicago, IL: University of Chicago Press (1963).
- Wever R. *The Circadian System of Man. Results of Experiments Under Temporal Isolation*. New York; Heidelberg: Springer-Verlag (1979). doi: 10.1007/978-1-4612-6142-1
- Tsaoausis I. Circadian preferences and personality traits: a meta-analysis. *Eur J Personal.* (2010) 24:356–73. doi: 10.1002/per.754

11. Keren H, Boyer P, Mort J, Eilam D. Pragmatic and idiosyncratic acts in human everyday routines: the counterpart of compulsive rituals. *Behav Brain Res.* (2010) **212**:90–5. doi: 10.1016/j.bbr.2010.03.051
12. Gill P, Arlitt M, Li Z, Mahanti A. Youtube traffic characterization: a view from the edge. In: *Proceedings of the 7th ACM SIGCOMM Conference on Internet Measurement*. San Diego, CA (2007). p. 15–28. doi: 10.1145/1298306.1298310
13. Kaltenbrunner A, Gómez V, Moghnieh A, Meza R, Blat J, López V. Homogeneous temporal activity patterns in a large online communication space. *IADIS Int J WWW/INTERNET* (2008) **6**:61–76.
14. Yasseri T, Sumi R, Kertész J. Circadian patterns of wikipedia editorial activity: a demographic analysis. *PLoS ONE* (2012) **7**:e30091. doi: 10.1371/journal.pone.0030091
15. Yasseri T, Quattrone G, Mashhadi A. Temporal analysis of activity patterns of editors in collaborative mapping project of OpenStreetMap. In: *Proceedings of the 9th International Symposium on Wikis and Open Collaboration*. Hong Kong (2013). p. 13. doi: 10.1145/2491055.2491068
16. Jo HH, Karsai M, Kertész J, Kaski K. Circadian pattern and burstiness in mobile phone communication. *New J Phys.* (2012) **14**:013055. doi: 10.1088/1367-2630/14/1/013055
17. Krings G, Karsai M, Bernhardsson S, Blondel VD, Saramäki J. Effects of time window size and placement on the structure of an aggregated communication network. *EPJ Data Sci.* (2012) **1**:1–16. doi: 10.1140/epjds4
18. Song C, Qu Z, Blumm N, Barabási AL. Limits of predictability in human mobility. *Science* (2010) **327**:1018–1021. doi: 10.1126/science.1177170
19. Ahas R, Aasa A, Silm S, Tiru M. Daily rhythms of suburban commuters' movements in the Tallinn metropolitan area: case study with mobile positioning data. *Transport Res Part C* (2010) **18**:45–54. doi: 10.1016/j.trc.2009.04.011
20. Louail T, Lenormand M, Cantu Ros OG, Picornell M, Herranz R, Frias-Martinez E, et al. From mobile phone data to the spatial structure of cities. *Sci Rep.* (2014) **6**:4. doi: 10.1038/srep05276
21. Aledavood T, López E, Roberts SGB, Reed-Tsochas F, Moro E, Dunbar RIM, et al. Daily rhythms in mobile telephone communication (2015). arXiv:1502.06866.
22. Saramäki, J, Moro, E. From seconds to months: an overview of multi-scale dynamics of mobile telephone calls. *Eur Phys J B.* (2015) **88**:164. doi: 10.1140/epjb/e2015-60106-6
23. Vajna S. *Modelling of Communication Dynamics*. Budapest University of Technology and Economics. Budapest (2012).
24. ten Thij M, Bhulai S, Kampstra P. Circadian patterns in Twitter. In: *DATA ANALYTICS 2014, The Third International Conference on Data Analytics*. Rome (2014) p. 12–17.
25. Noulas A, Scellato S, Mascolo C, Pontil M. An empirical study of geographic user activity patterns in foursquare. *ICWSM11* (2011).
26. Toole JL, Ulm M, González MC, Bauer D. Inferring land use from mobile phone activity. In: *Proceedings of the ACM SIGKDD International Workshop on Urban Computing*. Beijing: ACM (2012). p. 1–8. doi: 10.1145/2346496.2346498
27. Grauw S, Sobolevsky S, Moritz S, Gódor I, Ratti C. Towards a comparative science of cities: using mobile traffic records in New York, London, and Hong Kong. In: Hellich M, Arsanjani JJ, Leitner M editors. *Computational Approaches for Urban Environments*. Springer (2015). pp. 363–87. Available online at: <http://link.springer.com/book/10.1007/978-3-319-11469-9>
28. Dong Y, Pinelli F, Gkoufas Y, Nabi Z, Calabrese F, Chawla N. Inferring unusual crowd events from mobile phone call detail records. (2015). arXiv:1504.03643.
29. Riley M. *Iranian Hackers Stalked US Officials Report Says* (2014). Available online at: <http://www.bloomberg.com/news/articles/2014-05-29/iranian-hackers-stalked-u-s-officials-report-says>
30. Eckmann JP, Moses E, Sergi D. Entropy of dialogues creates coherent structures in e-mail traffic. *Proc Natl Acad Sci USA.* (2004) **101**: 14333–7. doi: 10.1073/pnas.0405728101
31. Eagle N, Pentland A. Reality mining: sensing complex social systems. *Pers Ubiquit Comput.* (2006) **10**:255–68. doi: 10.1007/s00779-005-0046-3
32. Karsai M, Kivela M, Pan RK, Kaski K, Kertész J, Barabási AL, et al. Small but slow world: how network topology and burstiness slow down spreading. *Phys Rev E.* (2011) **83**:025102. doi: 10.1103/PhysRevE.83.025102
33. Aledavood T, López E, Roberts SGB, Reed-Tsochas F, Moro E, Dunbar RIM, et al. Channel-specific daily patterns in mobile phone communication. (2015). arXiv:1507.04596.
34. Saramäki J, Leicht E, López E, Roberts SG, Reed-Tsochas F, Dunbar RI. Persistence of social signatures in human communication. *Proc Natl Acad Sci USA.* (2014) **111**:942–7. doi: 10.1073/pnas.1308540110
35. Refinetti R. *Circadian Physiology*. CRC Press (2005). doi: 10.1201/9781420039016
36. Ruffieux C, Marazzi A, Paccaud F. The circadian rhythm of the perinatal mortality rate in Switzerland. *Am J Epidemiol.* (1992) **135**:936–52.
37. Kouchaki M, Smith IH. The morning morality effect: the influence of time of day on unethical behavior. *Psychol Sci.* (2014) **25**:95–102. doi: 10.1177/0956797613498099
38. Blatter K, Cajochen C. Circadian rhythms in cognitive performance: methodological constraints, protocols, theoretical underpinnings. *Physiol Behav.* (2007) **90**:196–208. doi: 10.1016/j.physbeh.2006.09.009
39. Taillard J, Philip P, Bioulac B. Morningness/eveningness and the need for sleep. *J Sleep Res.* (1999) **8**:291–5. doi: 10.1046/j.1365-2869.1999.00176.x
40. Llorente A, García-Herranz M, Cebrian M, Moro E. Social media fingerprints of unemployment. *PLoS ONE* (2015) **10**:e0128692. doi: 10.1371/journal.pone.0128692
41. Deville P, Linard C, Martin S, Gilbert M, Stevens FR, Gaughan AE, et al. Dynamic population mapping using mobile phone data. *Proc Natl Acad Sci USA.* (2014) **111**:15888–93. doi: 10.1073/pnas.1408439111
42. Torous J, Staples P, Onnela JP. Tracking mental well-being: balancing rich sensing and patient needs. *Curr Psychiat Rep.* (2015) **17**:1–7. doi:10.1007/s11920-015-0602-0
43. Jain SH, Powers BW, Hawkins JB, Brownstein JS. The digital phenotype. *Nat Biotechnol.* (2015) **33**:462–3. doi: 10.1038/nbt.3223
44. Onnela JP. *Digital Phenotyping* (2015). Available online at: <http://www.hsph.harvard.edu/onnella-lab/research/>
45. Miller G. The smartphone psychology manifesto. *Perspect Psychol Sci.* (2012) **7**:221–37. doi: 10.1177/1745691612441215
46. Matthews M, Abdullah S, Gay G, Choudhury T. Tracking mental well-being: balancing rich sensing and patient needs. *Computer* (2014) **47**:36–43. doi: 10.1109/MC.2014.107
47. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, et al. Utilizing a personal smartphone custom app to assess the patient health questionnaire-9 (PHQ-9) depressive symptoms in patients with major depressive disorder. *JMIR Mental Health* (2015) **2**:e8. doi: 10.2196/mental.3889
48. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. *J Med Internet Res.* (2015) **17**:e175. doi: 10.2196/jmir.4273
49. Faurholt-Jepsen M, Frost M, Vinberg M, Christensen EM, Bardram JE, Kessing LV. Smartphone data as objective measures of bipolar disorder symptoms. *Psychiatry Res.* (2014) **217**:124–7. doi: 10.1016/j.psychres.2014.03.009
50. Barabási AL. The origin of burst and heavy tails in human dynamics. *Nature* (2005) **435**:207–11. doi: 10.1038/nature03459
51. Miritello G, Moro E, Lara R. Dynamical strength of social ties in information spreading. *Phys Rev E* (2011) **83**:045102. doi: 10.1103/PhysRevE.83.045102
52. Stopczynski A, Sekara V, Sapiezynski P, Cuttone A, Madsen MM, Larsen JE, et al. Measuring large-scale social networks with high resolution. *PLoS ONE* (2014) **9**:e95978. doi: 10.1371/journal.pone.0095978

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Aledavood, Lehmann and Saramäki. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



Emotions and Activity Profiles of Influential Users in Product Reviews Communities

Dorian Tanase, David Garcia, Antonios Garas and Frank Schweitzer*

Chair of Systems Design, ETH Zurich, Zurich, Switzerland

Viral marketing seeks to maximize the spread of a campaign through an online social network, often targeting influential nodes with high centrality. In this article, we analyze behavioral aspects of influential users in trust-based product reviews communities, quantifying emotional expression, helpfulness, and user activity level. We focus on two independent product review communities, *Dooyoo* and *Epinions*, in which users can write product reviews and define trust links to filter product recommendations. Following the patterns of social contagion processes, we measure user social influence by means of the *k*-shell decomposition of trust networks. For each of these users, we apply sentiment analysis to extract their extent of positive, negative, and neutral emotional expression. In addition, we quantify the level of feedback they received in their reviews, the length of their contributions, and their level of activity over their lifetime in the community. We find that users of both communities exhibit a large heterogeneity of social influence, and that helpfulness votes and age are significantly better predictors of the influence of an individual than sentiment. The most active of the analyzed communities shows a particular structure, in which the inner core of users is qualitatively different from its periphery in terms of a stronger positive and negative emotional expression. These results suggest that both objective and subjective aspects of reviews are relevant to the communication of subjective experience.

OPEN ACCESS

Edited by:

Taha Yasseri,
University of Oxford, UK

Reviewed by:

Boris Podobnik,
University of Rijeka, Croatia
Marija Mitrovic Dankulov,
Institute of Physics Belgrade, Serbia

*Correspondence:

David Garcia
dgarcia@ethz.ch

Specialty section:

This article was submitted to
Interdisciplinary Physics,
a section of the journal
Frontiers in Physics

Received: 26 June 2015

Accepted: 26 October 2015

Published: 17 November 2015

Citation:

Tanase D, Garcia D, Garas A and
Schweitzer F (2015) Emotions and
Activity Profiles of Influential Users in
Product Reviews Communities.
Front. Phys. 3:87.
doi: 10.3389/fphy.2015.00087

Keywords: social network analysis, social influence, sentiment, trust, spreading processes

1. INTRODUCTION

Popularity of socially-powered online platforms increased so much during the last years that, if we could imagine a country with a population as large as the user-base in Facebook, then it would be ranked as world's second largest country, with more than 1.23 Billion active users at the end of 2013 [1]. Users interact online via different platforms for personal blogging, dating, online shopping, reviewing products, etc. The latter two kind of platforms use their massive user community to both collect and disseminate information: Users create and discover reviews, form opinions based on the experience of others, and ultimately make the informed decision of buying a product or not. This form of socially-powered platforms are usually referred to as Social Recommender Systems (SRS) [2].

Similar to real-world social interactions, in online SRS platforms, some users manage to distinguish themselves from the rest by acquiring fame and social influence. If seen from a graph's perspective, some nodes become more central than others, but how this process works is not clear for real and online networks alike. How can a user increase its social influence and visibility?

Are there any similarities in the *career path* of successful users? In this article, we address these questions by performing an empirical analysis on two datasets of online SRS that contain both product reviews and explicit social networks. Information is transferred in these systems through social ties, by means of social recommender filtering, which selects products and reviews from the peers that a user trusts. This functionality creates a spreading process through the social network that offers opportunities for viral marketing [3], using the social capital of online communities to maximize the visibility of a product [4].

The emotional content in product reviews is an interesting resource not only to overcome the bias present in ratings, but for the role emotions play in human communication and product evaluation. Studies in social psychology show that people find emotional information more interesting than the non-emotional, and that they show more engagement with emotional narrators [5]. Additionally, the social link between narrator and listener has been observed to strengthen when emotions are involved [6]. We are interested in testing these social theories, and assess whether they hold also in online recommender systems: Does a user who shares its emotions have a larger impact in the community? Do users prefer neutral product evaluations or, on contrary, is the personal experience, as emotional as it can be, considered more valuable?

In the theory of core affect [7], emotions are partially conscious, short-lived internal states, as opposed to the nature of opinions. A reviewer might not be fully aware of its own emotions, and if asked a long time after making the review, these emotions would have relaxed or disappeared, while its opinion about a product would remain. There is an expected overlap between rating and emotional classification [8], but the properties and social dynamics of opinions and emotions differ. For example, disclosure of emotions has been shown to be a better predictor for social connection than the sharing of facts and information [9], and collective emotions pose additional questions regarding collective identity, social action, and emergent phenomena in human societies [10].

The topic of social influence and spreading processes in social networks has attracted increasing attention, due to the presence of frequent cascades and viral phenomena in social systems. Influence processes have been studied in the context of rumor spreading in social networks [11]. To identify social influence, traditional measures focused on the concept of centrality [12], often measuring it as degree or betweenness centrality [13]. Recent works have shown that coreness centrality [14, 15] outperforms degree and betweenness centrality in detecting influentials both data-driven simulations [16] leading to applications to political movements [17, 18], scientific rumors [19, 20], gender inequality in Wikipedia [21], and cascades of users leaving a social network [22].

Finding influentials is often motivated by viral marketing, aiming at the maximization of the reach of a marketing campaign and user adoption [4, 23, 24]. Beyond purchase decisions, users of social recommender systems create star ratings and write reviews that can influence product adoption. The straightforward manner to analyze these reviews is to take into account the star rating as a measure of consumer satisfaction. This approach

has been proved useful in the field of recommender systems [2, 25]. On the other hand, self-selection biases difficult the analysis of star-rating distributions, as their high bias reduces the heterogeneity of user evaluations, following a J-shaped distribution [26].

The large amount of product reviews in a social recommender system produce a state of information overload [25]. This kind of information overload influences the priority processing patterns of individuals [27]. Works in psychology identify emotions as one of the mechanisms for priority assignment: while we seek for positive experiences, negative ones make us react faster [28]. This leads to a stronger influence of emotions in social sharing [29], which also appears in product reviews [8]. Emotional expression cascades through social interaction have been identified in the context of chatrooms [30] and political movements [18], as well as for experimental [31] and field studies in social psychology [32]. Furthermore, pieces of information are more likely to be shared in a social context when they contain a stronger emotional content, as it has been shown for the case of urban legends [33].

Sentiment analysis tools allow researchers to process and analyze emotions in large scale datasets. Different techniques can be used to extract emotional content from short, informal texts [34, 35], being SentiStrength one of the leading tools for sentiment analysis in this context [36, 37]. Product reviews are much longer and better composed than tweets or YouTube comments, calling for the application of established lexicon-based techniques based on human annotation of words [35, 38]. These techniques have been proved useful to reveal patterns of depressive moods [39] and analyze the dynamics of happiness of whole societies [38]. We chose to apply this kind of lexicon-based sentiment analysis tool, due to its previous validation with large, formal texts, and for its possibility for extension to other languages [40].

To explore the role of emotions and activity into the social influence of users of product reviews communities, we empirically quantify user behavior in various aspects. First, we analyze the trust network of two independent online communities, measuring social influence in relation to spreading processes in social networks [41]. We compute the coreness centrality of all users [14], and validate that it serves as an indicator of the spreading potential of users. Second, we measure emotions in product reviews by means of sentiment analysis, and aggregated these values into emotional expression profiles of each users. Combining this subjective information with other objective dimensions, such as age in the community and review votes, we create extended user profiles with rich behavioral information. Third, we analyze the signatures of emotional expression across the different centrality values of each network, testing the existence of patterns of emotional expression.

2. MATERIALS AND METHODS

2.1. Product Reviews Communities Data

We base our empirical analysis on two independent datasets based on two trust-based product reviews communities:

Dooyoo¹ and Epinions². Dooyoo claims to be a “social-shopping platform which helps consumers make informed purchasing decisions”³. Similarly, Epinions is a product comparison website which features product reviews with a social component [42]. Both platforms are intended for English-speaking users, and allow them to post written reviews about products with a star-rating from 1 to 5. A particularly interesting feature of these two communities is that both allow the creation of directed social links that can be defined as trust and distrust links toward other users. Distrust links are not publicly available on the website, and for that reason our study is restricted only to trust links. These links are directional, meaning that the origin of the link trusts the destination of the link, as a way to acknowledge the quality of the reviews of the trusted user. The motivation for the creation of these links is advertised in both platforms as a way to improve product recommendations, as their recommender systems would refine the way they filter information based on this explicit trust [25].

Both platforms are product-generic, in the sense that users can review products in multiple categories, not limited to books or software. Apart from reviewing and creating trust links, users can also provide feedback about the quality of product reviews written by other users. This evaluation is done by clicking a helpful/unhelpful button, which the website uses to measure the helpfulness of a review as the aggregation of the votes of all users. This feedback feature is precisely relevant in Dooyoo, where users have the possibility of receiving money from the website as a reward for the creation of useful reviews⁴. In both communities, each review has a helpfulness score summarized as *Very helpful*, *Somewhat helpful*, *Helpful*, *Not helpful*, or *No feedback* if the review did not receive positive nor negative votes.

In our network datasets, nodes represent users, and a directed link from user u_1 to user u_2 means that u_1 explicitly trusts u_2 . In both communities, users are allowed to see all the reviews created by all the other users, i.e., there are no private reviews. This means that there is a global information flow between users, which does not necessarily depend on the trust network. On the other hand, both websites advertise that their recommender systems take into account trust links in order to personalize recommendations. This implies that the trust network exercises a “filtering influence,” increasing the visibility and impact of the reviews of user u_2 for user u_1 , if u_1 trusts u_2 . This opens the question of the role of the trust network, especially when users are allowed to see all the reviews and can vote any review, regardless of the trust network, as helpful or unhelpful.

For Dooyoo, we gather a dataset which we refer to as the DY dataset. Datasets on Epinions are available from previous work [42], but to the best of our knowledge, none of them used the text of the reviews for extracting additional information beyond ratings. Therefore, we performed a web crawl on Epinions and fetched, besides the trust network, the text of reviews. The raw data was further cleaned up, by removing duplicate reviews,

users, etc. We will refer to this dataset as the EP dataset. This second dataset is smaller, in terms of number of users, number of trust links and number of reviews than the version used in Walter et al. [25], but contains richer information including reviews text and helpfulness feedback. As shown in **Table 1**, the DY dataset contains roughly half the number of users in comparison to EP dataset, however, the amount of users that contributed at least one review is roughly the same. More details on the distributions of lifetimes and activity levels can be found in the Supplementary Information.

2.2. User Sentiment Analysis

The star-rating of a review provides the explicit opinion given by the user, but the emotional content is not acknowledged when making the review, contrary to other communities like Livejournal [43]. For this reason, we apply a sentiment analysis technique that extracts an estimation of the valence v , which represents the amount of pleasure or displeasure associated with an emotional experience [44]. Among other dimensions that can be used to measure emotions [45], valence is the one that explains the most variance of emotional experience [46, 47]. This technique analyzes each word in the review by looking into a lexicon on word valence, providing an estimation of v as the mean valence of the words appearing in the text (for more details see Supplementary Information). Then, this value of valence is compared with the baseline distribution of the valence for emotional words in generalized text, as estimated from a large dataset from web crawls [40]. If the valence of a review r is above a threshold given this baseline distribution, the review is classified as positive ($e_r = 1$), if it is below another threshold, it is classified as negative ($e_r = -1$), and if it is between both it is classified as neutral ($e_r = 0$).

Given the emotional classification of each review, we calculate the degree of positivity, negativity, and neutrality of every user, by aggregating its emotional scores over the whole number of reviews it contributed in the following way:

$$\begin{aligned} P_u &= \frac{1}{|R_u|} \sum_{r \in R_u} \Theta[e_r = 1] & N_u &= \frac{1}{|R_u|} \sum_{r \in R_u} \Theta[e_r = -1] \\ U_u &= \frac{1}{|R_u|} \sum_{r \in R_u} \Theta[e_r = 0] \end{aligned} \quad (1)$$

where R_u is the set of reviews written by the user u , $|R_u|$ is the number of reviews created by u , which is a metric for the

TABLE 1 | Descriptive statistics on the Dooyoo and Epinions datasets.

	Dooyoo (DY)	Epinions (EP)
Users	40,535	75,361
Reviews	524,000	101,595
Users with reviews	37,122	31,331
Trust links	199,061	544,469
Avg. rating	4.03	3.98
Avg. review length (words)	534	638
Avg. user lifetime (days)	100	138

¹<http://www.dooyoo.co.uk>.

²<http://www.epinions.com>.

³“About”-page of www.dooyoo.co.uk.

⁴Description of monetary rewards in Dooyoo: http://www.dooyoo.co.uk/community/_page/advice_participate.

amount of information it contributes to the community, and $\Theta(x)$ is a Boolean function that returns +1 if the argument is true and zero otherwise. These three metrics contain additional information about user behavior that is not contained in the average star-rating of a user.

Intuitively, one could expect that a successful user, a professional product reviewer, creates neutral, rigorous reviews, without emotional charge, in a similar fashion in which a journalist would write news and articles. However, in both datasets, we find that a large fraction of the reviews are positively charged, i.e., the user presents the product or service in a favorable manner by using positively emotional words. Reviews with negative emotions are less frequent than positive ones, but they are significantly present. These ratios are presented in **Table 2**.

2.3. Network Analysis

We quantify the social influence of users of Dooyoo and Epinions by analyzing their respective social networks. First, we measured a set of descriptive statistics on each network, measuring diameter, reciprocity, path length, and finding the largest weakly and strongly connected components. These metrics are included in **Table 3**, showing that a significant difference between the two datasets is the size of their largest strongly and weakly connected components. Beyond that difference, the rest of statistics show relative similarity, displaying typical properties of social networks such as low average path length and diameter. The reciprocity for both networks is relatively low, in line with previous findings on Twitter [48].

We measure the level of social influence of a user through the k-shell decomposition of the social network [14, 15, 18, 49]. We measure the influence of a node by its coreness centrality k_s , which is the state of the art metric to measure influence in

social networks, as it is the best known predictor for the size of cascades [16].

In general, the k-shell decomposition of a graph is obtained by recursively removing all its vertices with degree less than k , until all the remaining vertices have minimum degree $k + 1$. The removed vertices are labeled with a shell number (k_s) equal to k . For our study, we choose to collapse links into undirected ones, using as degree the sum of unidirectional and bidirectional links of a user. The reason for this stems from previous studies on Twitter, which show that the undirected k-shell decomposition of follower networks can predict empirical cascades of tweets in various phenomena [17, 50].

With the k-shell decomposition we are able to obtain a ranking of nodes which is related to a hierarchical organization in terms of importance, as illustrated in **Figure 1**. The larger the k_s of a node, the more influential it is. We should note that the coreness centrality is, in general, highly correlated with the degree centrality. However, there is no one to one relation, since as shown in **Figure 1**, a node can have large degree and still be located at an external shell. **Figure 2** shows the networks visualized with LaNet-vi [51], in which nodes have a color and position corresponding to their coreness.

3. RESULTS

3.1. Network Position and Social Influence

3.1.1. Heterogeneity of Coreness

For the EP network we find 126 shells, while for the DY network we find 84 shells. The distribution of coreness values k_s of both networks, shown in **Figure 3**, is skewed and reveals that the location of users in the k-shells follows similar patterns. The majority of users are located in the periphery of the network, and only a small fraction of them is placed in the more central k-shells. However, though, despite that the EP network is almost twice as large as the DY network (see **Table 1**; the LCC of the EP is more

TABLE 2 | The fraction of positively, negatively charged, and neutral reviews.

	Dooyoo UK	Epinions
Positive reviews	152,172 (29.05%)	24,104 (23.73%)
Negative reviews	93,649 (17.87%)	21,887 (21.54%)
Neutral reviews	278,096 (53.08%)	55,604 (54.73%)
Classifiable reviews	523,917 (99.98%)	101,595 (100%)

The percentages are calculated using the total number of classifiable reviews, because some reviews in the DY dataset were lacking emotion carrying words.

TABLE 3 | Network statistics of the analyzed datasets, Dooyoo and Epinions.

	Dooyoo (DY)	Epinions (EP)
Network diameter	13	16
Network reciprocity	0.2736	0.2151
Average path length	4.2299	4.8615
Largest SCC Size	10,200	30,396
Largest WCC Size	17,233	68,748

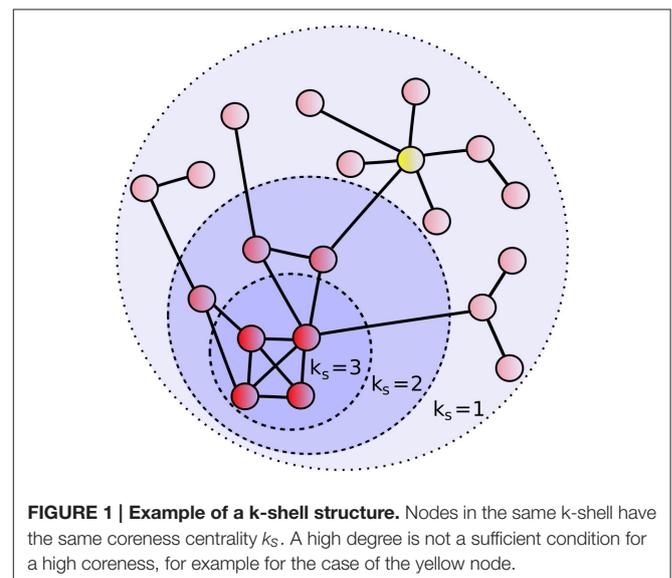
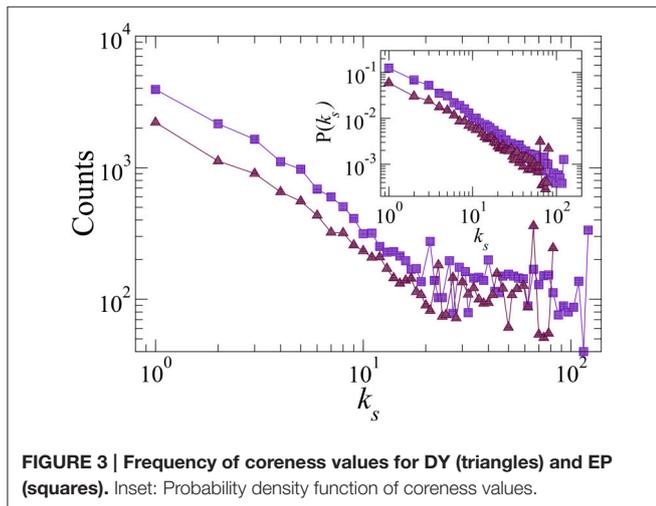
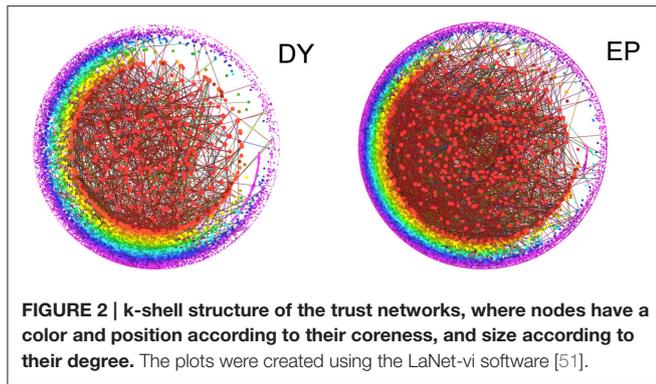


FIGURE 1 | Example of a k-shell structure. Nodes in the same k-shell have the same coreness centrality k_s . A high degree is not a sufficient condition for a high coreness, for example for the case of the yellow node.



than three times the LCC of the DY) the number of users in the more central k -shells is similar in both networks. This means that the number of very central users is not directly proportional to the total amount of users in a network, thus, there should be other factors determining users' centrality.

The heterogeneity of the distributions of k_s values becomes evident when fitting power-law distributions to the empirical data. Applying a maximum likelihood criterion that minimizes the Kolmogorov-Smirnov distance between empirical and theoretical distributions [52], we find that both distributions can be explained by truncated power laws of exponent $\alpha_{EP} = 1.39 \pm 0.004$ for EP and $\alpha_{DY} = 1.207 \pm 0.005$ for DY. This result is robust, since log-likelihood ratio tests vs. log-normal and exponential alternatives give positive and significant values, i.e., the power-law distribution explains the distribution of k_s significantly better than its non-scaling alternatives.

3.1.2. Social Influence Simulation

One of the goals of social networks is to facilitate information exchange between its users, i.e., information from user A can reach user B through the network link connecting them. Subsequently, the same piece of information can be forwarded by user B to user C through their respective link, and so on.

This is an example of a classical spreading process taking place in a network topology [41]. In product review communities an underlying explicit social network facilitates information exchange about products (i.e., reviews). For example, when a review is created, the peers of the author will get access to new information and they have the option to either read it (and become informed) or not. Therefore, a natural way to simulate information propagation in such systems is by means of a Susceptible-Infectious (or better suited to our case Susceptible-Informed, SI) model. Such models have been used widely in the literature to describe processes like the spreading of epidemics, rumors, economic crises, etc. [53–58].

We perform large scale computer simulations of spreading processes, assuming that users stay informed after reading a review, i.e., users do not return to the susceptible state. This SI process is modeled as follows: starting from the explicit social network (DY or EP) we choose a user at random and we assume it will try (through the creation of a review) to spread information to all users it is connected to. The probability that a targeted user becomes informed by reading the review is β , and remains constant throughout the simulation. Next, the informed users will try to pass this information to all their neighbors, and so on. This process is terminated after all informed users have tried to propagate information through their respective connections. For both networks, we perform 10 runs initiating the spreading process from a specific user, and we repeated this sequentially for every user in the network using probability of infection $\beta \in [0.1, 0.6]$ with step $\Delta\beta = 0.1$.

In **Figure 4**, we plot the average fraction f of users that become informed from reviews created by users belonging to a k -shell vs. the k -shell number (k_s). In agreement with [16], we find that information initiated by the more central users in terms of k_s can reach a larger percentage of users in both networks. Therefore, the incentive of increasing ones impact in the network is correlated with the network centrality. As a result if users want to increase the impact of the transmitted information, they should try to become more central.

In the left panel of **Figure 5** we plot the average fraction, f_c , of the network that becomes informed by a review created from users belonging to the Largest Connected Component (LCC) of the network vs. the probability of transmission β . Besides the expected trend that f_c increases with the probability β , in the left panel of **Figure 5** it is shown that in the DY network f_c can receive much higher values for the same β than in the EP network. This result suggests that the DY network allows a more efficient information transmission in comparison to the EP network, if we only consider the Largest Connected Component (LCC). But, if we consider the full network, then the situation is inverted. This can be attributed to the different connectivity pattern observed in the two communities (as discussed in **Table 1**), where for EP the largest connected component is almost 90% of the nodes, while for DY this percentage is almost 40%.

We calculate topological features of users measured through the k -shell decomposition neglecting any possible effect of directionality in the links that connect them. However, the evolution of a dynamical process on a network could be heavily affected by the presence of directed links. Thus, in order to test

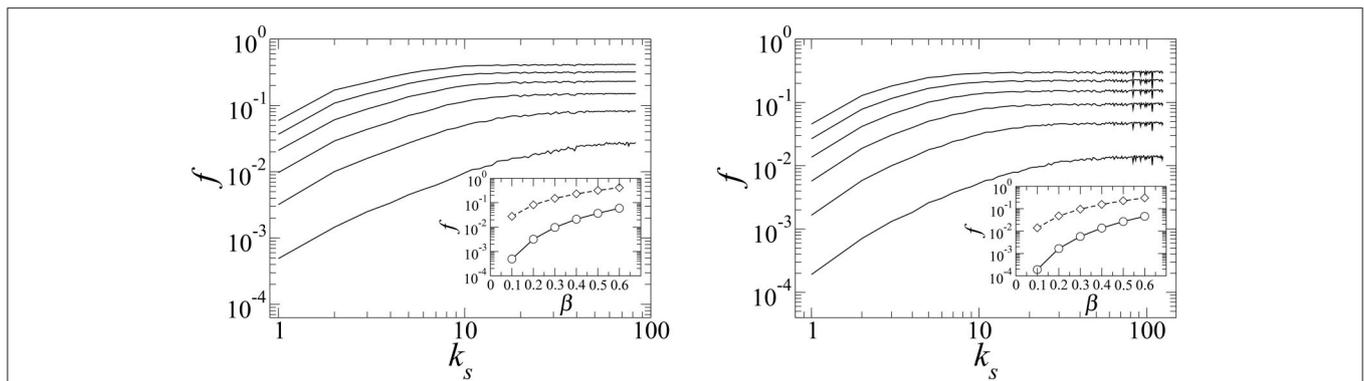


FIGURE 4 | Average fraction f of informed population by reviews created from users of different k_s vs. the k -shell number (k_s). The different curves show results for different probabilities $\beta \in [0.1, 0.6]$, with the lower curves corresponding to smaller β 's. Inset: Average fraction f of informed population by reviews created from users of the lowest (circle) and the highest (diamond) k -shell vs. probability of information transmission β . **Left panel: DY. Right panel: EP.**

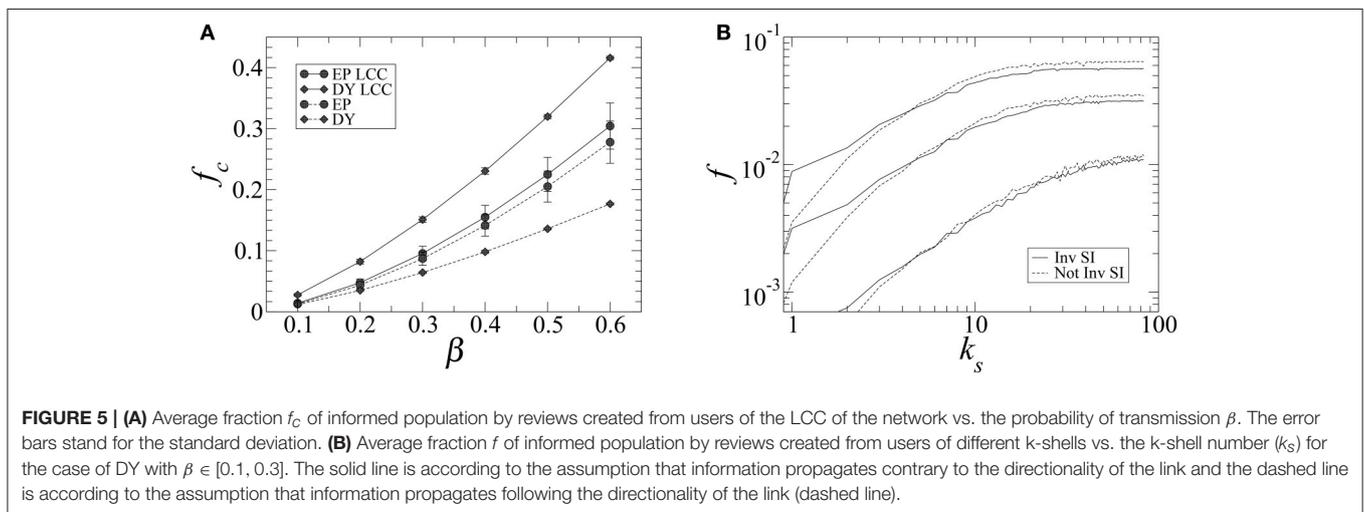


FIGURE 5 | (A) Average fraction f_c of informed population by reviews created from users of the LCC of the network vs. the probability of transmission β . The error bars stand for the standard deviation. **(B)** Average fraction f of informed population by reviews created from users of different k -shells vs. the k -shell number (k_s) for the case of DY with $\beta \in [0.1, 0.3]$. The solid line is according to the assumption that information propagates contrary to the directionality of the link and the dashed line is according to the assumption that information propagates following the directionality of the link (dashed line).

whether link directionality affects our conclusions we apply the SI model to the DY network assuming two distinct hypotheses, (a) that information flows according to the direction of the links, and (b) that information flows inversely to the direction of the links. The right panel of **Figure 5** shows the fraction f vs. k_s for both hypotheses described above i.e., information flows following the link directionality, and information flows in the opposite direction. In general, we find that for $k_s > 5$ the link directionality does not influence heavily the process of spreading, thus, the results we discussed in the previous analysis are valid for both cases. In what follows we try to identify the profile of the more central users, in order to understand whether there are common patterns in their behavior. After all, it is natural to assume that they did not end up being central purely by “luck.”

3.2. User Production

3.2.1. Helpfulness

Users give feedback on the quality of other users’ reviews by voting individual reviews as helpful or unhelpful. In both

TABLE 4 | Ratios of community feedback values for the reviews of each dataset.

	Dooyoo UK	Epinions
Very helpful	354,927 (67.73%)	48,575 (47.81%)
Somewhat helpful	22,190 (4.23%)	17,023 (16.76%)
Helpful	137,750 (26.29%)	35,340 (34.79%)
Not helpful	1540 (0.29%)	1 (0.00%)
No feedback	7593 (1.45%)	656 (0.65%)

communities, each review has a *helpfulness rating* calculated as a combination of these votes. The helpful rating h_r is displayed along with a review r in a qualitative scale of four grades: “very useful,” “useful,” “somewhat useful,” and “not useful.” We map these ratings on a scale from 0 (not useful) to 4 (very useful), in order to quantify the impact of a review in the community. **Table 4** contains the ratios of each type of feedback in EP and DY.

Given this measure of helpfulness of a review, for each user u we can calculate a value of total helpfulness

$$h_u = \sum_{r \in R_u} h_r \tag{2}$$

which is a sum of all the helpfulness scores attributed by the community to the reviews created by the user, R_u . **Figure 6** shows the distribution of the values of h_u in each community. This figure reveals the large heterogeneity in the helpfulness of users, where most users have very few helpful reviews, while some others accumulate large amounts of positive feedback from the rest. The two communities differ in the shape of this heterogeneity, as in DY there are significantly larger amounts of users with high helpfulness in comparison with EP.

While the distribution of h_u in EP is very irregular, it seems to follow a stylized broad distribution in DY. While the tail is not long enough to verify a power-law distribution [59], we tested the possibility of a log-normal distribution. A maximum likelihood estimation, discussed in the Supplementary Information, gives a set of parameters that fail to fit the tail of the distribution, leading us to reject the log-normal hypothesis. This initial observation indicates the existence of a process of helpfulness accumulation that creates larger heterogeneity than the one present in a log-normal distribution, but we do not have enough data to precisely explore its properties at larger scales.

3.2.2. Ratings and Emotions

Product reviews contain factual information about properties of the product and its experienced quality from the reviewer’s point of view. In the two communities we study, as discussed above, a product review contains two elements: a star rating, which summarizes product experience in a form of opinion, and a review text with detailed information written by the user. The straightforward manner to analyze these reviews is to take into account the star rating, as a measure of consumer satisfaction with the product. This approach has been proved useful in the field of recommender systems [2, 25, 60, 61]. On the other hand, self-selection biases make it difficult to analyze star-rating distributions, as their high bias reduces the heterogeneity of user evaluations, following a J-shaped distribution [26]. This is the case for both EP and DY, where the distribution of star-ratings of

the reviews follows a J-shaped distribution, as shown in **Figure 7**. Most of the reviews have star ratings ≥ 4 , with a small increase on the amount of 1-star reviews in comparison with 2-star reviews. In addition, user average ratings suffer from this bias, as shown in **Figure 8**. To overcome this limitation, we study the emotions expressed in the text of the review, as explained below.

Figure 8 shows the scatter plots of the user ratios of emotional expression vs. the average rating of users, with the corresponding distributions in each axis. We can clearly observe how the average rating of users, r_u is skewed with a mean around 4, while the ratios N_u , U_u , and P_u have different distributions between 0 and 1. The pairwise Pearson correlation coefficients of r_u with each of the other three variables has absolute values below 0.25, indicating that there is significant variance of the emotional expression of users that is not captured by the ratings. The three metrics N_u , U_u , and P_u provide us with additional data beyond the simple average rating provided by a user, profiling the different types of users by the way they express their emotions in the reviews they create.

3.3. The Profile of Influential Users

We test whether there are user specific features associated with an increased coreness of the user k_u and thus with an increased user social influence. For our analysis, we use a linear regression technique on a logarithmic transformation of k_u , using the behavior metrics explained above as independent variables. This technique of substitution models has been used before to study the relation between Facebook user popularity and personality metrics from a survey [62]. In our case, we fit the following model:

$$\log(k_u + 1) = \alpha + \beta_P P_u + \beta_N N_u + \beta_R r_u + \beta_T \log(t_u) + \beta_H \log(h_u + 1) + \beta_W \log(w_u) \tag{3}$$

The dependent variable is a transformation of the coreness in two ways: (i) calculating the logarithm to provide a monotonic transformation that decreases the variance of k_u , as its distribution is right skewed (see **Figure 2**), and (ii) an increment of 1 to include in our analysis active but disconnected nodes

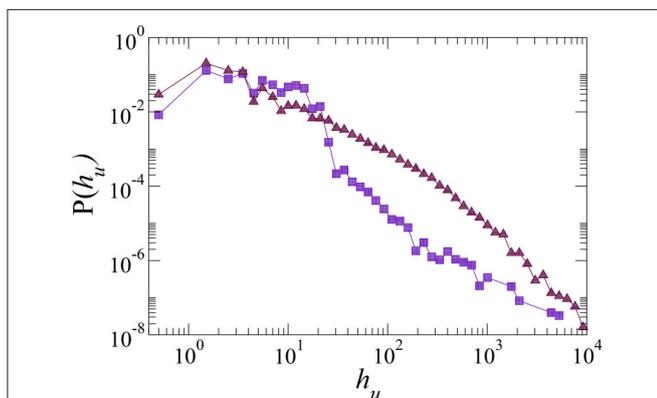


FIGURE 6 | The distribution of the total helpfulness (h_u) of users for DY (triangles) and EP (squares).

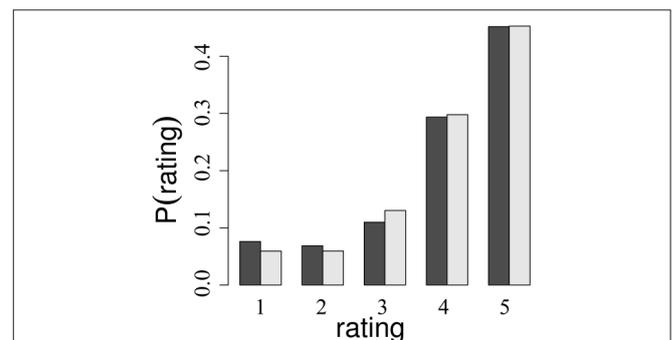
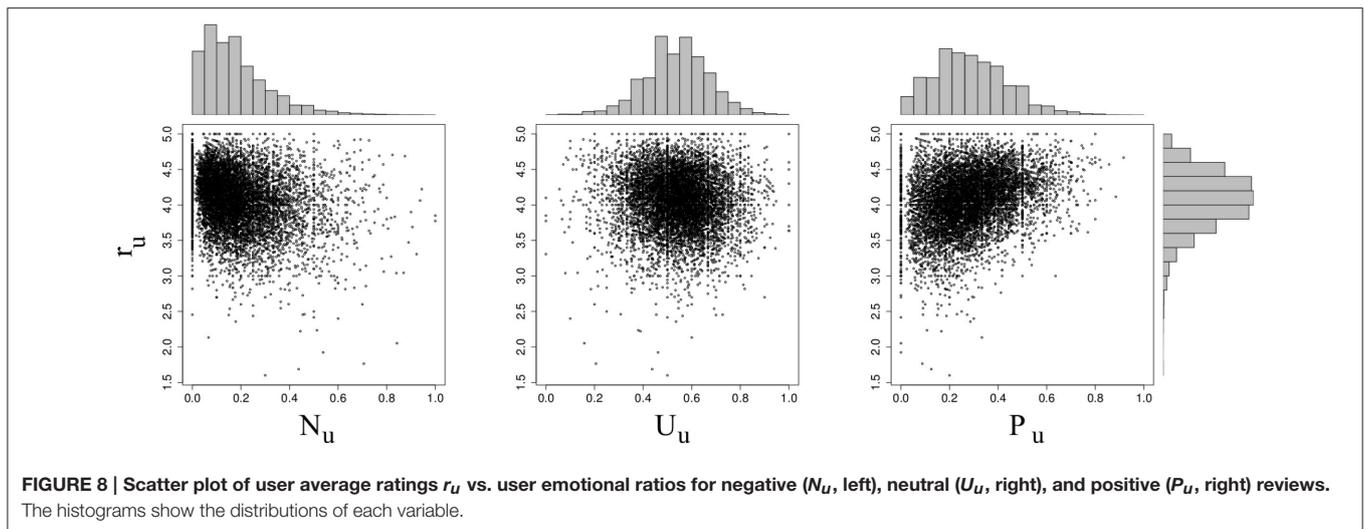


FIGURE 7 | Distribution of ratings in the reviews of EP (dark) and DY (light). Both distributions show a strong bias toward positive ratings, with a moderated J-shape.



with $k_u = 0$. The independent variables of our model capture the different metrics of user behavior explained above. The first two variables, P_u and N_u account for the emotional expression of the user. We omit the ratio of neutral messages U_u , as its redundancy with the previous two would lead to a singularity due to the identity $P_u + N_u + U_u = 1$. The third variable, the average rating of the user r_u accounts for the style of the user in capturing its opinions into a precise number. The fourth variable is the lifetime of the user in the community t_u , as explained in Section 3.2.1. This variable accounts for heterogeneity in the age of users, and it might play a relevant role in the impact a user can have in the product reviews community. The fifth variable is a transformation of the total helpfulness of the user h_u , following the same principle as for the dependent variable. Finally, the last variable accounts for the logarithm of the average amount of words in the reviews of the user $\log(w_u)$, as a proxy for the amount of unfiltered information in a typical review of the user, which could have an effect on its relevance in the community (for more details on the amount of words of reviews, see SI).

We fit Equation 3, first normalizing each variable and then solving the linear regression by the method of least squares, obtaining results summarized in Table 5. Our first observation is that the linear regression is different for the two datasets. The R^2 for the case of DY is 0.6174, while for EP is 0.1751. This indicates that the data we obtained for D○○○○ allows us to better estimate the social influence of a user by its activity, in comparison with the EP dataset. Second, in both cases the largest significant coefficient is the total helpfulness of the user. This shows that the total helpfulness and the k-shell number of a user are directly related. In other words, a user becomes central, and therefore, more important in the community, if it contributes with many helpful reviews.

The second largest weight for the users in DY corresponds to the lifetime of a user in the community t_u , with significant positive value. This means that users that have been longer in the product reviews community also have higher coreness. For EP, the average length of the reviews created by a user is the second

TABLE 5 | Linear regression coefficients and p -values for $\log(k_u + 1)$ from the rest of the user metrics (normalized), for D○○○○ (DY) and Epinions (EP).

	DY	EP
Variable	Weight	Weight
Intercept	0.782***	1.527***
$\log(w_u)$	0.072***	0.259***
$\log(h_u + 1)$	0.805***	0.360***
$\log(t_u)$	0.121***	-0.032***
P_u	-0.008*	0.047***
N_u	0.009*	0.007
r_u	0.022***	0.061***

Significance levels: * $p < 0.1$, *** $p < 0.001$.

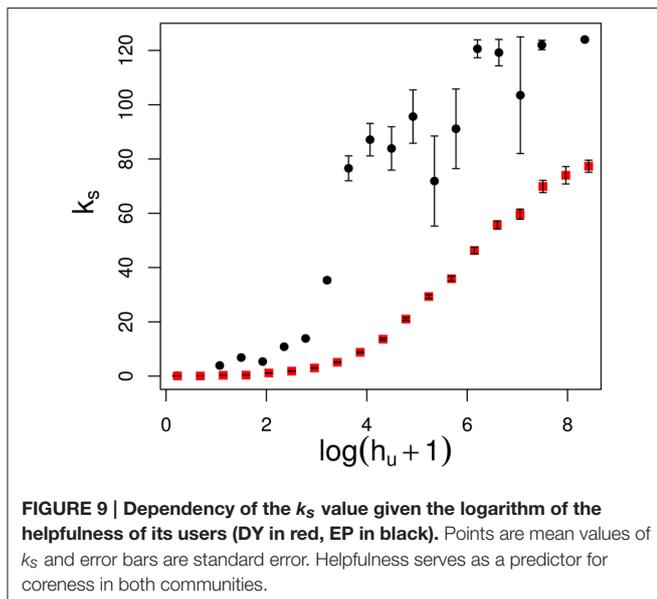
most important factor for centrality. As in DY with lifetime, w_u is less relevant than the total helpfulness implying that the community is not concerned about the size of reviews but rather about their overall quality.

Focusing on the relation between the coreness of a user and its total helpfulness, we computed Pearson’s correlation coefficients between $\log(h_u + 1)$ and $\log(k_u + 1)$, giving a value of 0.677 ± 0.006 for DY, and 0.337 ± 0.01 for EP, both with $p < 0.001$. This way, we conclude that the total helpfulness of a user is a good predictor for its network centrality, as both variables are significantly correlated in both datasets. Figure 9 shows the mean coreness values for users of different helpfulness levels. Both communities display a clear relation between both variables: users with higher amounts of helpful reviews also have more social influence.

Testing the role of emotionality ratios and average rating in the results of Table 5, we notice that all three variables have very low regression weights. P_u and N_u have low significance in DY, and N_u is not significant in EP. This indicates that the role of emotions in social influence cannot be observed through this analysis at the individual level, and that helpfulness and age are more predictive variables.

3.4. The Emotional Core of Dooyoo

Motivated by the theory of collective emotions [10], we tackle the question of how do the aggregated emotions of users in different k-shells differ. For a given coreness number k_s , we aggregate the activity of all the users in that shell by the average values $\langle P \rangle_s$, $\langle U \rangle_s$, $\langle N \rangle_s$, calculated over all the users with coreness k_s . The emotional profile of the users in different k-shells can be observed in **Figure 10**, where each k-shell is represented by a semicircle with distance to the center according to their coreness number. Each shell has three colors that range from the minimum to the maximum values of each $\langle N \rangle_s$, $\langle U \rangle_s$, $\langle P \rangle_s$. For both communities, k-shells closer to the core have stronger negativity and weaker neutrality. It is important to notice that, even though these emotions increase within their individual ranges, the maximum values of $\langle N \rangle_s$ in DY still remain lower than the other two average ratios.



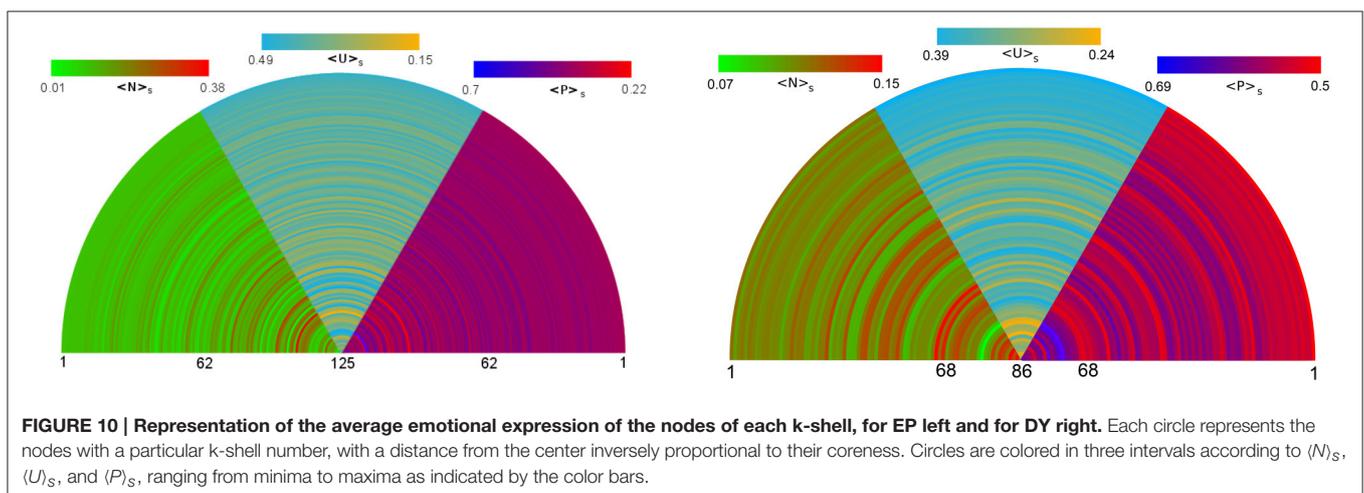
A close inspection of **Figure 10** shows a pattern in DY that does not appear in EP: There is an inner core composed of some shells with high coreness number that have stronger average emotion indicators, as compared with the rest of shells with lower k_s numbers. This inner vs. outer part difference is described by a critical value of k_c , which highlights a stronger emotional expression for k-shells with k_s at least k_c (the core), in comparison with the weaker emotional expression of those with $k_s < k_c$ (the periphery).

We test the existence of this core by a set of Wilcoxon tests dividing each community in users with k-shell number above and below different values of k_c . **Figure 11** shows the Wilcoxon distances Δ of $\langle N \rangle_s$, $\langle U \rangle_s$, and $\langle P \rangle_s$ between the core and periphery, for values of the division k_c from 1 to the maximum coreness number. For EP we did not find any significant nonzero distances separating the neutral and negative average scores of the inner and outer parts. For DY, on the other hand, the scenario is different. There is a value $k_c = 68$, where there is a sharp transition that indicates a maximal distinction between core and periphery, highlighting the existence of a more emotional central subcommunity.

The significant separation of DY in core and periphery leads to a central core with stronger emotional expression. The right panel of **Figure 11** shows the Wilcoxon distance between emotion ratios, comparing core and periphery divided by $k_c = 68$. The core has significantly higher negative and positive ratios, with decreased neutrality ratio. This result is supported by the dependence of the p -value of the Wilcoxon test and the ratios of emotional expression vs. k_c , as shown in the SI.

4. DISCUSSION

Our analysis of two online product reviews communities shows the relation between community feedback, emotions, and social influence within the trust network. We measure social influence by means of the coreness of individual users, and validated such metric based on the SI process of information spread. Our findings show that, in line with previous research [16],



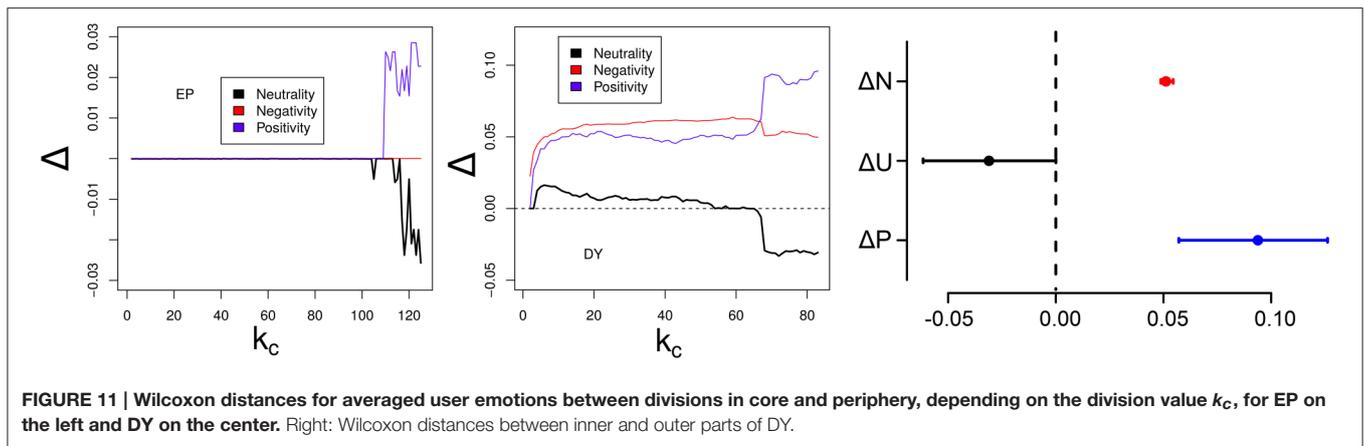


FIGURE 11 | Wilcoxon distances for averaged user emotions between divisions in core and periphery, depending on the division value k_c , for EP on the left and DY on the center. Right: Wilcoxon distances between inner and outer parts of DY.

the expected size of a cascade increases with the coreness centrality of the node it starts from. Furthermore, we analyze the heterogeneity of coreness through model fitting to the empirical distributions, finding that the coreness in both communities follows a power-law distribution. The exponents we found for these fits suggest that the mean and variance of coreness scales with system size, i.e., larger online communities serve as training grounds for even more influential users. Testing this type of scaling requires the analysis of several online communities, and remains open for future research.

We measure emotional expression in reviews through the ANEW lexicon, and aggregate the emotions of individual users in three scores for positivity, negativity, and neutrality. These three dimensions create a richer representation of individuals beyond average ratings, as emotional expression contains information not encoded in the star-ratings of reviews. Combining these features with the lifetime in the community, the average review size in words, and the levels of helpfulness votes of the users, we find that total helpfulness and average review length are the most relevant indicators for individual social influence, beyond emotional expression. Our observational analysis of one snapshot of the system point at the relevance of emotions in social influence, but further research should test other individual and temporal aspects of this explanation. Experimental studies can isolate the individual components that drive the decisions and expressions of users. Data with temporal resolution in network formation should further explore the career path of influential users, measuring the changes in k -core values as a function of contributions and emotions.

Our statistical analysis shows the existence of a sharp transition in coreness that divides the Dooyoo community in two levels: An emotional core and a more neutral surface. This structure was absent in Epinions, opening the question what process could create such difference in the relation between topology and emotional expression. An initial conjecture would point to the different reward schemes of the two communities:

Dooyoo offered monetary rewards to its most successful users, who created the emotional core of influential users. While our results at the individual level are inconclusive with respect to emotional expression, this characterization of emotions in a core-periphery structure suggests that the expression of emotions provides a medium for the communication of subjective experience. Such kind of communication process would enhance the interaction of certain types of users, improving their social influence as a whole rather than if they just wrote reviews with purely factual information. Understanding how such a pattern emerges from individual emotional interaction is a question open for future research, which could potentially link individual and collective patterns of emotions and social influence.

AUTHOR CONTRIBUTIONS

DT gathered and processed data, DT and AG analyzed the networks, DT and DG performed statistical analyses, DT, DG, AG, and FS wrote the article.

FUNDING

This research has received funding from the European Community's Seventh Framework Programme FP7-ICT-2008-3 under grant agreement no 231323 (CYBEREMOTIONS).

ACKNOWLEDGMENTS

The authors would like to thank Epinions.com and Dooyoo.co.uk for their accessibility to public reviews and trust data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fphy.2015.00087>

REFERENCES

1. Facebook. *Facebook Reports Fourth Quarter and Full Year 2013 Results* (2014). Available online at: <http://investor.fb.com/releasedetail.cfm?ReleaseID=821954>
2. Victor P, Cornelis C, De Cock M. *Trust Networks for Recommender Systems*. Vol. 4. Springer Science & Business Media (2011). Available online at: <http://www.springer.com/us/book/9789491216077>
3. Leskovec J, Adamic LA, Huberman BA. The dynamics of viral marketing. *ACM Trans Web (TWEB)* (2007) 1:5. doi: 10.1145/1232722.1232727
4. Bonchi F, Castillo C, Gionis A, Jaimes A. Social network analysis and mining for business applications. *ACM Trans Intell Syst Technol.* (2011) 2:22. doi: 10.1145/1961189.1961194
5. Collins NL, Miller LC. Self-disclosure and liking: a meta-analytic review. *Psychol Bull.* (1994) 116:457-475.
6. Peters K, Kashima Y. From social talk to social action: shaping the social triad with emotion sharing. *J Pers Soc Psychol.* (2007) 93:780–97. doi: 10.1037/0022-3514.93.5.780
7. Russell JA. Core affect and the psychological construction of emotion. *Psychol Rev.* (2003) 110:145. doi: 10.1037/0033-295X.110.1.145
8. Garcia D, Schweitzer F. Emotions in product reviews - empirics and models. In: *International Conference on Social Computing*. Boston, MA (2011). pp. 483–8.
9. Laurenceau JB, Barrett LF, Pietromonaco PR. Intimacy as an interpersonal process: the importance of self-disclosure, partner disclosure, and perceived partner responsiveness in interpersonal exchanges. *J Pers Soc Psychol.* (1998) 74:1238–51. doi: 10.1037/0022-3514.74.5.1238
10. von Scheve C, Salmela M. *Collective Emotions*. Oxford, UK: Oxford University Press (2014).
11. Chierichetti F, Lattanzi S, Panconesi A. Rumor spreading in social networks. *Theor Comput Sci.* (2011) 412:2602–10. doi: 10.1016/j.tcs.2010.11.001
12. Newman MEJ. *Networks: An Introduction*. Oxford University Press (2010). Available online at: <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780199206650.001.0001/acprof-9780199206650>
13. Freeman LC. A set of measures of centrality based on betweenness. *Sociometry* (1977) 40:35–41. doi: 10.2307/3033543
14. Seidman SB. Network structure and minimum degree. *Soc Netw.* (1983) 5:269–87. doi: 10.1016/0378-8733(83)90028-X
15. Garas A, Schweitzer F, Havlin S. A k-shell decomposition method for weighted networks. *New J Phys.* (2012) 14:083030. doi: 10.1088/1367-2630/14/8/083030
16. Kitsak M, Gallos LK, Havlin S, Liljeros F, Muchnik L, Stanley HE, et al. Identification of influential spreaders in complex networks. *Nat Phys.* (2010) 6:888–93. doi: 10.1038/nphys1746
17. González-Bailón S, Borge-Holthoefer J, Moreno Y. Broadcasters and hidden influentials in online protest diffusion. *Am Behav Sci.* (2013) 57:943–65. doi: 10.1177/0002764213479371
18. Alvarez R, Garcia D, Moreno Y, Schweitzer F. Sentiment cascades in the 15M movement. *EPJ Data Sci.* (2015) 4:1–13. doi: 10.1140/epjds/s13688-015-0042-4
19. De Domenico M, Lima A, Mougél P, Musolesi M. The anatomy of a scientific rumor. *Sci Rep.* (2013) 3:02980. doi: 10.1038/srep02980
20. Borge-Holthoefer J, Meloni S, Gonçalves B, Moreno Y. Emergence of influential spreaders in modified rumor models. *J Stat Phys.* (2013) 151:383–93. doi: 10.1007/s10955-012-0595-6
21. Wagner C, Garcia D, Jadidi M, Strohmaier M. It's a man's wikipedia? Assessing gender inequality in an online encyclopedia. In: *International AAAI Conference on Weblogs and Social Media*. Oxford, UK (2015). pp. 454–63.
22. Garcia D, Mavrodiev P, Schweitzer F. Social resilience in online communities: the autopsy of Friendster. In: *Conference in Online Social Networks*. Boston, MA (2013). pp. 39–50.
23. Arthur D, Motwani R, Sharma A, Xu Y. Pricing strategies for viral marketing on social networks. In: Leonardi S, editor. *Internet and Network Economics*. Rome: Springer (2009). pp. 101–12.
24. Barbieri N, Bonchi F. Influence maximization with viral product design. In: *Proceedings of the 2014 SIAM International Conference on Data Mining* (2014). pp. 55–63. Available online at: <http://epubs.siam.org/doi/abs/10.1137/1.9781611973440.7>
25. Walter FE, Battiston S, Schweitzer F. Personalised and dynamic trust in social networks. In: *Proceedings of the Third ACM Conference on Recommender Systems - RecSys '09*. New York, NY (2009). p. 197.
26. Hu N, Zhang J, Pavlou PA. Overcoming the J-shaped distribution of product reviews. In: *Communications of the ACM* (2009) 52:144. Available online at: http://dl.acm.org/ft_gateway.cfm?id=1562800&type=html
27. Rodriguez MG, Gummadi K, Schölkopf B. Quantifying information overload in social media and its impact on social contagions. In: *Proceedings Eighth International AAAI Conference on Weblogs and Social Media (ICWSM)*. Ann Arbor, MI (2014).
28. Norman GJ, Norris CJ, Gollan J, Ito Ta, Hawkey LC, Larsen JT, et al. Current emotion research in psychophysiology: the neurobiology of evaluative bivalence. *Emotion Rev.* (2011) 3:349–59. doi: 10.1177/1754073911402403
29. Rime B. Emotion elicits the social sharing of emotion: theory and empirical review. *Emotion Rev.* (2009) 1:60–85. doi: 10.1177/1754073908097189
30. Garas A, Garcia D, Skowron M, Schweitzer F. Emotional persistence in online chatting communities. *Sci Rep.* (2012) 2:402. doi: 10.1038/srep00402
31. Christophe V, Rime B. Exposure to the social sharing of emotion: emotional impact, listener responses and secondary social sharing. *Eur J Soc Psychol.* (1997) 27:37–54.
32. Harber KD. The emotional broadcaster theory of social sharing. *J Lang Soc Psychol.* (2005) 24:382–400. doi: 10.1177/0261927X05281426
33. Heath C, Bell C, Steinberg E. Emotional selection in memes: the case of urban legends. *J Pers.* (2001) 81:1028–41. doi: 10.1037/0022-3514.81.6.1028
34. Gonçalves P, Araújo M, Benevenuto F, Cha M. Comparing and combining sentiment analysis methods. In: *Conference on Online Social Networks*. Boston, MA (2013).
35. Iosub D, Laniado D, Castillo C, Fuster Morell M, Kaltenbrunner A. Emotions under discussion: gender, status and communication in online collaboration. *PLoS ONE.* (2014) 9:e104880. doi: 10.1371/journal.pone.0104880
36. Thelwall M, Buckley K, Paltoglou G, Skowron M, Garcia D, Gobron S, et al. Damping sentiment analysis in online communication: discussions, monologs and dialogs. In: Gelbukh A, editor. *Computational Linguistics and Intelligent Text Processing*. Samos: Springer (2013). pp. 1–12.
37. Abbasi A, Hassan A, Dhar M. Benchmarking Twitter sentiment analysis tools. In: *Language Resources and Evaluation Conference*. Reykjavik (2014).
38. Dodds PS, Danforth CM. Measuring the happiness of large-scale written expression: songs, blogs, and presidents. *J Happiness Stud.* (2009) 11:441–56. doi: 10.1007/s10902-009-9150-9
39. Park M, Cha C, Cha M. Depressive moods of users portrayed in twitter. In: *Proceedings of the ACM SIGKDD Workshop on Healthcare Informatics, HI-KDD*. Beijing (2012).
40. Garcia D, Garas A, Schweitzer F. Positive words carry less information than negative words. *EPJ Data Sci.* (2012) 1:3. doi: 10.1140/epjds3
41. Anderson RM, May RM. *Infectious Diseases of Humans: Dynamics and Control*. Oxford, UK: Oxford Science Publications (1992).
42. Massa P, Avesani P. Controversial users demand local trust metrics: an experimental study on opinions. com community. In: *Proceedings of the National Conference on Artificial Intelligence*. Vol. 20. Pittsburgh, PA (2005). p. 121.
43. Zafarani R, Cole WD, Liu H. Sentiment propagation in social networks : a case study in livejournal. In: *Advances in Social Computing, Lecture Notes in Computer Science*. Vol. 6007. Bethesda, MD (2010). pp. 413–20.
44. Russell JA. A circumplex model of affect. *J Pers Soc Psychol.* (1980) 39:1161–78. doi: 10.1037/h0077714
45. Bradley MM, Lang PJ. *Affective Norms for English Words (ANEW): Instruction Manual and Affective Ratings*. The Center for Research in Psychophysiology, University of Florida (1999).
46. Russell JA, Mehrabian A. Evidence for a three-factor theory of emotions. *J Res Pers.* (1977) 11:273–94.
47. Fontaine JR, Scherer KR, Roesch EB, Ellsworth PC. The world of emotions is not two-dimensional. *Psychol Sci.* (2007) 18:1050–7. doi: 10.1111/j.1467-9280.2007.02024.x

48. Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? In: *International Conference on World Wide Web*. Raleigh, NC (2010). pp. 591–600.
49. Batagelj V, Zaveršnik M. Fast algorithms for determining (generalized) core groups in social networks. *Adv Data Anal Class.* (2011) **5**:129–45. doi: 10.1007/s11634-010-0079-y
50. Baños RA, Borge-Holthoefer J, Moreno Y. The role of hidden influentials in the diffusion of online information cascades. *EPJ Data Sci.* (2013) **2**:1–16. doi: 10.1140/epjds18
51. Beiró MG, Alvarez-Hamelin JJ, Busch JR. A low complexity visualization tool that helps to perform complex systems analysis. *New J Phys.* (2008) **10**:125003. doi: 10.1088/1367-2630/10/12/125003
52. Alstott J, Bullmore E, Plenz D. powerlaw: a Python package for analysis of heavy-tailed distributions. *PLoS ONE.* (2014) **9**:e85777. doi: 10.1371/journal.pone.0085777
53. Daley DJ, Kendall DG. Stochastic rumours. *IMA J Appl Math.* (1965) **1**:42–55. doi: 10.1093/imamat/1.1.42
54. Hethcote HW. The mathematics of infectious diseases. *SIAM Rev.* (2000) **42**:599. doi: 10.1137/S0036144500371907
55. Newman MEJ. Spread of epidemic disease on networks. *Phys Rev E.* (2002) **66**:1–11. doi: 10.1103/PhysRevE.66.016128
56. Colizza V, Barrat A, Barthélemy M, Vespignani A. The role of the airline transportation network in the prediction and predictability of global epidemics. *Proc Natl Acad Sci USA.* (2006) **103**:2015–20. doi: 10.1073/pnas.0510525103
57. Castellano C, Fortunato S, Loreto V. Statistical physics of social dynamics. *Rev Mod Phys.* (2009) **81**:591–646. doi: 10.1103/RevModPhys.81.591
58. Garas A, Argyrakis P, Rozenblat C, Tomassini M, Havlin S. Worldwide spreading of economic crisis. *New J Phys.* (2010) **12**:113043. doi: 10.1088/1367-2630/12/11/113043
59. Clauset A, Shalizi CR, Newman MEJ. Power-law distributions in empirical data. *SIAM Rev.* (2009) **51**:661. doi: 10.1137/070710111
60. Jamali M, Ester M. A matrix factorization technique with trust propagation for recommendation in social networks. In: *Proceedings of the Fourth ACM Conference on Recommender Systems - RecSys '10*. New York, NY: ACM Press (2010). p. 135. Available online at: <http://portal.acm.org/citation.cfm?doid=1864708.1864736>
61. Tanase D. *Social Dynamics in Online Product Review Communities*. Dissertation, Eidgenössische Technische Hochschule ETH Zürich, Zürich (2013). Available online at: <http://e-collection.library.ethz.ch/view/eth:6433>
62. Quercia D, Lambiotte R, Stillwell D, Kosinski M, Crowcroft J. The personality of popular facebook users. In: *Conference on Computer Supported Cooperative Work*. Seattle, WA (2012).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Tanase, Garcia, Garas and Schweitzer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.



A Biased Review of Biases in Twitter Studies on Political Collective Action

Peter Cihon¹ and Taha Yasseri^{2,3*}

¹ Williams College, Williamstown, MA, USA, ² Oxford Internet Institute, University of Oxford, Oxford, UK, ³ Alan Turing Institute, London, UK

OPEN ACCESS

Edited by:

Matjaž Perc,
University of Maribor, Slovenia

Reviewed by:

Renaud Lambiotte,
Université de Namur, Belgium
Kunal Bhattacharya,
Birla Institute of Technology and
Science, India
Cornelius Puschmann,
Alexander von Humboldt Institute for
Internet and Society, Germany

*Correspondence:

Taha Yasseri
taha.yasseri@oii.ox.ac.uk

Specialty section:

This article was submitted to
Interdisciplinary Physics,
a section of the journal
Frontiers in Physics

Received: 17 May 2016

Accepted: 26 July 2016

Published: 09 August 2016

Citation:

Cihon P and Yasseri T (2016) A Biased
Review of Biases in Twitter Studies on
Political Collective Action.
Front. Phys. 4:34.
doi: 10.3389/fphy.2016.00034

In recent years researchers have gravitated to Twitter and other social media platforms as fertile ground for empirical analysis of social phenomena. Social media provides researchers access to trace data of interactions and discourse that once went unrecorded in the offline world. Researchers have sought to use these data to explain social phenomena both particular to social media and applicable to the broader social world. This paper offers a minireview of Twitter-based research on political crowd behavior. This literature offers insight into particular social phenomena on Twitter, but often fails to use standardized methods that permit interpretation beyond individual studies. Moreover, the literature fails to ground methodologies and results in social or political theory, divorcing empirical research from the theory needed to interpret it. Rather, investigations focus primarily on methodological innovations for social media analyses, but these too often fail to sufficiently demonstrate the validity of such methodologies. This minireview considers a small number of selected papers; we analyse their (often lack of) theoretical approaches, review their methodological innovations, and offer suggestions as to the relevance of their results for political scientists and sociologists.

Keywords: social media, twitter, mobilization, campaign, collective action, bias, theory

1. INTRODUCTION

Since its founding in 2006, Twitter has become an important platform for news, politics, culture, and more across the globe [1]. Twitter, like other social media platforms, empowers new forms of social organization that were once impossible. Margetts et al. discuss changing conceptions of membership and organization on social media [2]; Twitter communities and conversations need not be bounded by geography, propinquity, or social hierarchy. As a result, social and political movements have taken to the site as a means of organizing activity both online and offline. In facilitating these movements, Twitter simultaneously makes available a data trail never before seen in social research. Researchers have embraced these data to create an expanding body of literature on Twitter and social media writ large. On the other hand some researchers have been more skeptical about using social media data in general, and specially data from Twitter, in studying social behavior [3]. And some others question the relevance of such data to social sciences completely; see **Figure 1** for a satirical illustration of this view.

This literature is quite diverse. Some investigations seek to relate Twitter to the offline world [4]. Kwak et al. [5] crawl the entirety of Twitter and find that the platform's social networks differ from offline sociability in important ways. Huberman et al. [6] examine user behavior in addition to network structure, and find strong "friend" relationships, akin to offline sociability, are important predictors of Twitter activity. Gonçalves et al. [7] use Twitter data to validate anthropologist Robin

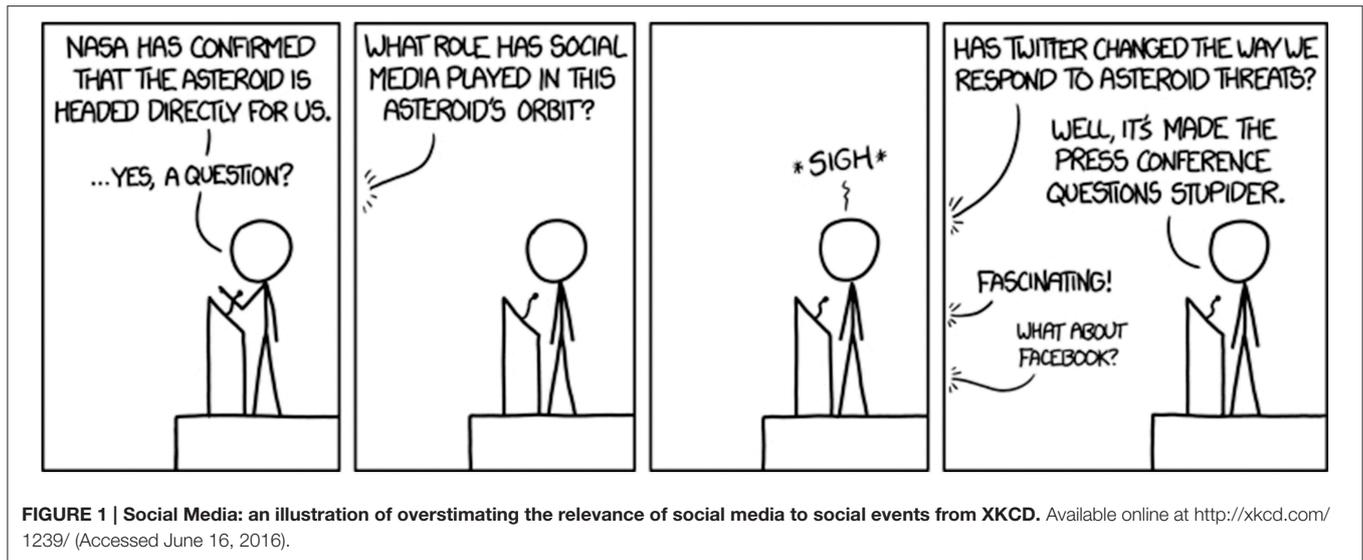


FIGURE 1 | Social Media: an illustration of overestimating the relevance of social media to social events from XKCD. Available online at <http://xkcd.com/1239/> (Accessed June 16, 2016).

Dunbar's proposed quantitative limit to social relationships. Still other investigations analyze the various uses of Twitter. Examining social influence, Bakshy et al. [8] study Twitter cascades, and find that the largest are started by past influential users with many followers. Semantic investigations in various languages and national contexts have been quite popular [9–11]. Questions of how Twitter and platform phenomena map onto offline geographic have also been widely studied [12, 13].

Yet, this body of literature is only unified in the source of its data; it remains fractured across many disciplines and fails to establish set procedures for drawing conclusions from these rich datasets; for an earlier survey of the literature see Jungherr [14]. Indeed, metareviews of election prediction using Twitter have raised significant concerns of this literature's validity [15, 16]. This minireview extends this critical discussion of Twitter literature to political action. We selected the reviewed studies in order to sample a variety of topics and methodologies, however this collection is not exhaustive by any means and hence we named the paper a "biased review." The approach we have taken in this work deviates from systematic reviews in the field such as ones described in Petticrew and Roberts [17]. Our sample purposively draws a geographic diversity of papers studying Twitter-based political action in Europe, the Middle East, and the United States. Yet, some gaps certainly remain, including the glaring absence of hashtag activism studies and terrorist propaganda activity, two topics important to political action on Twitter that warrant further study. Hence we acknowledge that our review is not inclusive in terms of coverage of all the relevant papers in the field. For a general overview of studies on online behavior see DiMaggio et al. [18].

In reviewing the state of Twitter literature on political action, we seek to explain the role of computational social science (also called social data science) methodologies in augmenting political scientific and sociological understanding of these phenomena. Our minireview is structured as follows. We begin by examining the role of theory, and find that most often authors do not consider the expansive political and social theoretical literature in

their analyses of online social phenomena. Instead, they provide case studies and methodological developments exclusively for Twitter research. We next examine the methodologies of these studies, and, drawing upon Ruths and Pfeffer [3], we find that many papers fail to support their choice of methodology within the greater literature. We then examine significant results and discuss implications for further Twitter studies of political action.

2. WHERE IS THE THEORY?

Social and political theory serves an important role in making sense of social research by fitting individual studies into larger theoretical frameworks. In this way, individual studies can intelligibly inform future research. Alternatively, data analysis without a coherent, defensible theoretical framework serves only to explain a single observation at one point in time. The papers reviewed here fall into three broad categories in their use of theory: no theory, theory-light, and theory-heavy. Papers fall into these categories irrespective of methodological or phenomenological focus.

Papers without theoretical grounding may cursorily cite but fail to engage theoretical texts. Beguerisse-Díaz et al. [19] examine communities and functional roles on Twitter during the UK riots of August 2011. To explain these phenomena, however, they cite no social theory. While the authors offer sophisticated methodical innovations for determining interest communities and individual roles in those communities, they do so without reference to a broader social science literature. Some other investigations offer cursory theory in their discussions of Twitter data. Borge-Holthoefer et al. [20] investigate political polarization surrounding the events that precipitated Egyptian President Morsi's removal from power in 2013. Their analysis of changes in loudness of opposing factions, although quite enlightening, is not grounded in any theoretical model of political action. Instead, the authors proceed based on a number of platform-specific assumptions that do not readily permit results

to be generalized beyond Twitter. The authors suggest their findings contribute to the study of bipolar societies yet do not develop a theoretical model for such applications. The authors do use social theory, however sparingly, in order to contextualize their results, but even here theoretical discussion is lacking. Conover et al. [21] study partisan communities and behavior on Twitter during the 2010 U.S. midterm elections. They similarly prioritize analytical innovations over theoretical explanation. The authors analyze behavior, communication, and connectivity between users, but do not seek to explain observed partisan differences. Their research yields statistically significant differences between liberal and conservative communities in follower and retweet networks, which begs the question: why do these differences exist? Such explanations could benefit from examining elections literature to develop a general theoretical model of partisan sharing. Although the authors do briefly address the 2008 U.S. presidential election, it is only to contrast resulting phenomena, not to offer explanatory theories.

In contrast, Alvarez et al. [22] explain political action in the Spanish 15M movement using Durkheimian theory of collective identity and establish their work on firm basis in collective action literature. Yet, while the authors base their methodology in theory, their findings do not directly engage with that theory aside from “quantifying” it. A similar fate befalls sampled predictive studies, which draw on theory to produce empirical results, but often fail to engage those results with underlying theory. Weng et al. [23] develop a model that predicts viral memes using community structure, based on theoretical insight from contagion theory. The authors find that viral memes spread by simple contagion, in contrast to unsuccessful memes which spread via complex contagion; still, only the briefest theoretical discussion for this result is offered. Garcia-Herranz et al. [24] develop a methodological innovation using individual Twitter users as sensors for contagious outbreaks based in the “friendship paradox” and contagion theory. This mechanism uses network topology as an effective predictor, but does not address the social phenomena that create and sustain that topology. Such methodological innovations provide researchers new analytical tools for observational analysis, but these tools remain of dubious explanatory value because they fail to ground methods in theory of the social world.

Twitter data present an opportunity not simply for analysis of social interactions on the platform but, if done well, these insights hold potential to contribute to new visions of the social world. Rigorous data science can generate new theory. Coppock et al. [25] are particularly notable in this regard. The authors base their methodological innovation in Twitter mobilization inducement on an extensive theoretical literature review, which yields three opposing hypotheses. They assess the political theory of collective action as it applies to Twitter via these three hypotheses, and find that the Civic Voluntarism Model is most consistent with their results. Likewise, González-Bailón et al. [26], in their study of protest recruitment dynamics in the Spanish 15M movement, offer both an extensive grounding in social theory and theory-engaging results. The authors’ findings serve to clarify threshold models of political action and “collective effervescence.”

As to the particular theories addressed, the above mentioned papers focus primarily on political action and network theories of diffusion and contagion. Important in such topics, but absent from all investigations, is discussion of power or hierarchy. Although Twitter may permit communication between the powerful and powerless, it does not do so in a vacuum. The platform operates within numerous contexts, e.g., the offline influence of particular users and the online influence of those with numerous followers. Reconciling methodologies with theories of power promises to provide further insight into political action on Twitter. More broadly, a greater focus on theory is needed for Twitter analyses to provide externally valid insight into the social world, both online and off.

3. DIVERGENCE IN METHODS

In developing analyses of Twitter data, researchers have not drawn on a coherent body of agreed-upon methodologies. Rather, methodological choices differ considerably from one paper to another. Ruths and Pfeffer [3] offers a critique of many common social media analysis practices. Drawing from that work as well as our own insights, we examine many of the methodological choices made in our sample papers. We have delineated these choices into several overarching categories: data, filtering, networks and centrality, cascades and communities, experiments, and conjecture.

Before addressing the methodological choices outlined above, we first address several important findings from Ruths and Pfeffer [3]. Today, academic research writ large—including social media work and much more—is insufficiently transparent. Academic journals publish only “successful” studies. Without publishing methodologies that failed to explain political action phenomena, how is one to weigh the probability that the supposed “fit” observed is not due to random chance? Even those papers which address the robustness of their analysis, often stop at a very shallow significance tests using p -value, which is argued to be a flawed practice [27, 28]. Similarly, when new methodologies are created, as in Weng et al. [23], Garcia-Herranz et al. [24] and Coppock et al. [25], they are justified vis-à-vis random baselines and not prior methods. New methods are useful, but are they better than existing tools? These opacity critiques are fundamental to the current state of Twitter scholarship. Researchers should be cognizant of these limitations when drawing conclusions from their work and should alter their methodologies to account for these limitations whenever possible.

3.1. Data

Twitter data ultimately comes from the Twitter platform. If scholars wish to make claims about the versatility of their methodologies and findings, they must justify their data-collection methods as representative of underlying populations—on Twitter or elsewhere. This proves a problematic task. The Twitter API offers researchers an incredible array of tweet, user, and more data for analysis; yet, the API acts as a “blackbox” filter that may not yield representative data [29, 30]. For example, Weng et al. [23] “randomly” collect 10% of public

tweets for one month from the API. Not only does the API preclude analysis as to the representativeness of the sample but it too prevents researchers from comparing studies over time, as the API sampling algorithm itself will change. Proprietary sampling methods only further exacerbate the opacity problem. In González-Bailón et al. [26], the authors use a proprietary sampling method to generate their dataset of Spanish tweets from Spain. The authors of Borge-Holthoefer et al. [20] do as well, using Twitter4J¹ and TweetMogaz² as data sources.

Other papers do not use a global sampling method, but obtain data in other ways. Beguerisse-Díaz et al. [19] use a list of “influential Twitter users” published in *The Guardian* as the starting point for data collection. Coppock et al. [25] develop their experimental design in cooperation with the League of Conservation Voters, and use their Twitter followers as test subjects. Other papers, including Conover et al. [21] and Alvarez et al. [22] collect data by following particular hashtags and the users who tweeted them. Garcia-Herranz et al. [24] collect Twitter data by snowball-sampling from one influential user, Paris Hilton, as well as all users mentioning trending topics. None of these sampling methods allows authors to make broad claims about the Twitter platform and political action in general. The method used in Garcia-Herranz et al. [24] is particularly concerning, as it attempts to collect a large sample to sufficiently model a Twitter population, but the choice of method undermines this very goal.

A final complication of data in Twitter studies regards the publication of that data. Once data is collected and analyzed, it is rarely made available for others to replicate these studies—the hallmark of good research. The problem here lies with Twitter itself; the terms of use preclude the republication of tweet contents that have been scraped from the site³.

3.2. Filtering

Following data collection, researchers often filter an intractable dataset into a manageable sample. Researchers often use filtering to select a coherent sample. Language and geography offer clear examples. Borge-Holthoefer et al. [20] limit their dataset to Arabic tweets about Egypt. Both González-Bailón et al. [26] and Alvarez et al. [22] limit their datasets to Spanish tweets from Spain. To do so, however, both papers use a proprietary filtering process from Cierzo Development Ltd⁴. As addressed above, proprietary methodologies stymie research transparency and replication.

Filtering can likewise facilitate a narrowing of research focus given a particular sample population. One common means of achieving a relevant dataset is to use hashtags as labels for tweets in which they appear. In González-Bailón et al. [26] the authors obtain a sample of protest-related tweets using a list of 70 hashtags affiliated with the Spanish 15M movement. Conover et al. [21] filter to a sample of political tweets using a list of political hashtags and, in an excellent technique, allow the list

of hashtags to grow based on co-occurring hashtags. In Borge-Holthoefer et al. [20] the authors go one step further, and query not only hashtags but complete tweet content. Arabic tweets were normalized for spelling and filtered by a series of Boolean queries with a set of 112 relevant keywords.

Researchers, after filtering for a relevant sample and topic, may further filter for user attributes. Borge-Holthoefer et al. [20] restrict their sample to high activity users with more than ten tweets extant in the limited sample. Beguerisse-Díaz et al. [19] limit their dataset to users central in the friend-follower network, those in the giant component. Users outside the giant component generally had incomplete Twitter information, and, as such, were dropped from the analysis. Weng et al. [23] limits the data to only reciprocal relationships. Conover et al. [21] filter tweets with geo-tags. The authors use a self-reported location field as their data source, despite the fact that someone can put “the moon” or anything else as their location. Indeed, Graham et al. use linguistic analyses to determine that such user-provided locations are poor proxies for true physical location [13]. Although the authors acknowledge the preliminary status of their analysis and its utility as an illustration of potential data-driven hypotheses, it left us unsatisfied with a lack of methodological rigor that should underlie even the most tentative of filtering claims.

Authors may choose to filter for no other reason than to obtain a manageable dataset. Such decisions need not be arbitrary. Garcia-Herranz et al. [24] settle on a particular sample size for their analyses, seeking to balance statistical power and the need to keep test and control groups from overlapping in the network. The authors offer an effective defense of their decision, presenting brief analyses of other sample sizes as well. Coppock et al. [25], on the other hand, arbitrarily remove Twitter users with more than 5000 followers from their sample because, they argue, these users are “more likely” to be influential or organizations, and therefore differ from the rest of the sample. This decision to remove outliers and the arbitrariness of the choice of threshold introduces systematic biases in the results, fundamentally undermining their analyses.

These myriad filtering decisions often go insufficiently defended. Those who do defend filtering choices often do so without referencing past literature. Even sound filtering decisions, however, undermine the general claims researchers can make. This may be one reason most of the studies fail to contribute to social theory beyond their micro case studies.

3.3. Networks and Centrality

Twitter lends itself to fruitful network analyses—of both explicit interactions and other derivative relations. Conover et al. [21] use three network projections to analyze partisan political behavior during the 2010 U.S. midterm elections: one network sees users connected when mentioned together in a tweet, another where users are linked by retweeting behavior and third, the original explicit user follow-ship network. Weng et al. [23] also uses three networks—mention, retweet, and follow—to study meme virality. The authors conduct primary analysis on the follow network and use the other two as robustness tests. In studying protest recruitment to the 15M movement, González-Bailón et al. [26] make use of two networks, one symmetric (comprised

¹<http://twitter4j.org/en/index.html>

²<http://www.tweetmogaz.com/>

³Twitter Terms of Service: <https://twitter.com/tos?lang=en>

⁴Formerly <http://www.cierzo-development.com>; see archive at <http://tinyurl.com/jzbewt8>

of reciprocated following relationships) and one asymmetric to study protest recruitment to the 15M movement. The authors use these networks to determine the influence of broadcasting users. Still other authors use single, traditional follower networks in their analyses [24, 25].

Network analyses are all the more powerful when they are combined, as in Weng et al. [23] and González-Bailón et al. [26]. In Borge-Holthoefer et al. [20], the authors offer another insight when they use network analyses over time with temporally evolving networks in response to events that preceded Egyptian President Morsi's removal from power. The authors recreate a sequence of networks that evolve over time. This method offers insight into how online activity responds to offline events in Egypt, and could be a powerful tool in many other contexts, helping to parse a key question of political action: how groups respond to events and evolve over time. The opposite, to assume a network remains static during a given period, precludes this insight and undermines social analysis. In González-Bailón et al. [26], the authors exemplify this pitfall, as a network of protesters being recruited surely saw significant changes during their study's time period. Given the fast growing literature on temporal networks [31, 32], more attention is required in analyzing the dynamics of networked political activities.

Beyond decisions of network type and temporality, authors make important choices in projecting and using Twitter networks. Weng et al. [23] does not weight network edges based on number of tweets, and chooses to limit the network projection to reciprocal relationships. Both decisions fundamentally affect results, and undermine its validity as representing activity on the Twitter platform. Others, including González-Bailón et al. [26] account for asymmetry in their network projections.

In doing network analysis, many researchers use centrality scores as a means to find the most influential users. Researchers have developed a number of different definitions and algorithms for centrality [33]. The choice of a specific approach, however, depends on the particular context and research questions. Often times this choice is not well justified in the given context of online political mobilizations. Among the papers considered here, *k*-core centrality [34] is the most common choice [21, 22, 26]. While *k*-core centrality is a very useful tool to find the backbone of the network, it neglects social brokers, or the nodes with high betweenness centrality,—relevant features in their own right when studying social behavior [35].

3.4. Cascades and Communities

Whether in networks or another form, Twitter data yield insight through a multitude of different analytical techniques. One such technique examines tweets as they flow through the network in cascades. Cascades follow a single tweet that is retweeted or similar tweets as they move across a network. The Twitter platform makes these analyses difficult, however, as retweets are connected to the original tweet, not the tweet that triggered the retweet [3]. Researchers address this pitfall by using temporal sequencing to order and connect tweets or retweets. To achieve meaningful results, studies must sufficiently filter the tweets to establish that sequential tweets are related in content as well as

time, which undermines representativeness, as discussed above [20, 22, 23].

Another common technique examines tweet content. Alvarez et al. [22] analyze their data for its social and sentiment content using semantic and sentiment analytic algorithms that analyzes tweets based on a test set. The authors use this technique to draw conclusions of individual users' opinions of the 15M movement in Spain by analyzing up to 200 authored tweets on the topic per user. This technique holds great promise for future studies of political activity, and indeed any activity, on Twitter. Borge-Holthoefer et al. [20] use a less sophisticated solution toward a similar goal: they characterize users as either for or against military intervention in Egypt. The authors attempt to show changes in opinion, and so cannot not rely on comprehensive opinion from a mass of past tweets as done in Alvarez et al. [22]. Instead, Borge-Holthoefer et al. [20] uses coded hashtags to indicate users' opinions. Although this technique allows for discernable changes in opinion, the authors establish a dichotomy that threatens to oversimplify users' opinions.

Community detection is another key analytical tool for Twitter researchers. Using network topology or node (user or tweet) content, researchers can cluster similar nodes and provide insight into social systems on a macroscopic scale. There are a variety of techniques, each with its own set of strengths and weaknesses. Weng et al. [23] uses the Infomap algorithm [36] and test the robustness of their results by applying a second community detection technique, Link Clustering. Conover et al. [21] uses a combination of two techniques, Rhagavan's label propagation method [37] seeded with node labels from Newman's leading eigenvector modularity maximization [38]. The authors selected this combination of methods because it "neatly divides the population ... into two distinct communities." Yet, the authors fail to defend these observations rigorously in their paper. Beguerisse-Díaz et al. [19], on the other hand, effectively defend their decisions in setting resolution parameters for the Markov Stability method [39]. The authors also use community detection creatively in conjunction with a functional role-determining algorithm to assign "roles" to users without a priori assignments of those groups. Borge-Holthoefer et al. [20] select an apt community detection method that corresponds well with their objectives: to follow changes in polarity over time, the authors use label propagation, whereby nodes spread their assigned polarity. This method allows for seeding with nodes of known belief—useful in monitoring the progression of the Egyptian protests on Twitter, as many important actors' positions were publicly known. Yet this decision too comes with a cost: the authors program the label propagation to allow for only two polarities: Secularist or Islamist, even though they acknowledge that a third camp likely existed, namely supporters of deposed Hosni Mubarak.

While community detection is still considered as an open question in network science, both at the definition and algorithmic implementation levels [40], many papers use one or more of these methods without enough care to make sure that the methods and definitions that they are using in their specific problem is well justified.

3.5. Experiments

Twitter also lends itself as an experimental platform for researchers to implement controlled studies of social phenomena. In particular, Garcia-Herranz et al. [24] and Weng et al. [23] seek to predict viral memes on Twitter using network topology and activation in linked users and communities, respectively. Coppock et al. [25] run two experiments on inducing political behavior on Twitter using different types and phrasing of messages. In all cases, authors necessarily use controls in their experimental context. Garcia-Herranz et al. [24] create a null distribution of tweets with randomly shuffled timestamps to distinguish the effect of user centrality from user tweeting rate. Weng et al. [23] use two baseline models to quantify the predictive power of their community-based model. The authors use a random guess and community-blind predictor, against both of which the model is highly statistically significant. Coppock et al. [25], with a true experimental design, offer an extensive discussion of experimental controls on Twitter. The platform has inherent limitations for public tweet experiments because there is no effective way to separate experimental and control users given an inherently interconnected network structure. But the authors design their study to use direct messages to selected users as the experimental variable. The authors even tweaked and repeated the study to improve randomization in the control. Such a methodology makes [25] an example of a particularly strong experimental Twitter paper.

3.6. Conjecture

As we have seen, Twitter provides researchers myriad analytical techniques. Methodological choices as to which techniques to use present a fundamental challenge for researchers. They must select and properly defend their choice of methods that both work and fit their theoretical objectives. As we have noted above, there are numerous instances where researchers will do better jobs than others are achieving a methodological fit and defending it in their studies. Some researchers may face the temptation to extend analyses to produce exciting results, but do so at the expense of sound methodologies. Future Twitter research would be well served to stress defensible, rigorous methodologies that are couched within existing theoretical literature from the social sciences, something that is rare today.

4. WHAT DID WE LEARN?

Taken collectively, the reviewed investigations offer considerable insight into political activities conducted on the Twitter platform, through analyses that examine political action in the abstract and others that offer case studies of concrete political action. These insights particularly address the roles of communities and individual users, connections between such entities, as well as the content they tweet. Predictive models take these insights and offer tools for, perhaps, understanding political action in real-time. Garcia-Herranz et al. use a sensor group of central users to predict virality of content, and extend this predictive sensor beyond Twitter to Google searches [24]. Weng et al. use connection topology to predict virality, although the predictive model is not extended to other content [23]. González-Bailón

et al. observe viral tweets emerge from randomly distributed seed users, indicating exogenous factors determine the origins of viral content [26]. Taken together, these three studies offer an understanding of mass communication on Twitter: viral content tends to originate randomly across the platform, reach more central users first, and spread across communities more easily than non-viral content. Theoretical explanations of what makes viral content in the first place, however, is lacking in these analyses, and warrants further attention.

Given a methodological focus, topology can offer insights into its embedded users. Beguerisse-Díaz et al. [19] use topographical analyses to reveal flow based roles, interest communities, and individual vantage points without a priori assignment. Conover et al. [21] assign political leaning and then examine differences in partisan topologies in communities, tweeting activity, retweeting behavior, and mentions. Both approaches offer insight into political behavior using topology, with different strengths. The techniques used in Beguerisse-Díaz et al. [19] are quite useful when the partisan landscape on a particular issue is unknown; The approach in Conover et al. [21] yields greater understanding of known divisions.

Topology is not the sole determinant of activity, however, and tweet content analyses offer a second means of understanding political activity on Twitter. Alvarez et al. [22] finds that, in the context of the Spanish 15M indignados, tweets with high social and negative content spread in larger cascades. Tweet content also readily lends itself to analyses which link Twitter with offline phenomena. Borge-Holthoefer et al. [20] and González-Bailón et al. [26] find that, in 2013 Egyptian protests and Spanish 15M protests, respectively, real world events impact tweeting behavior. Coppock et al. [25] successfully induce off-Twitter behavior using the content of tweets. Content analyses offer insight into non-platform-dependent political activity.

Topology and content are distinct analyses. Research that combines the two to answer a single question can yield robust results. Several papers attempt this, Borge-Holthoefer et al. [20] most successfully. The authors use content analysis to classify tweets and users into opinion groups, and then create temporally based retweet networks to follow changes in the activity and composition of those opinion groups. Alvarez et al. [22] use content analysis of observed network topological phenomena, e.g., cascades, to quantify the social and emotional effects of content on sharing outcomes. Beguerisse-Díaz et al. [19] too combine methodologies, although less rigorously: they use word clouds to label topologically derived network communities.

In this vein, many of the above mentioned investigations could benefit from incorporating mixed methodologies and drawing on each others analyses. Future research should seek to emulate the approach in Borge-Holthoefer et al. [20]. Further use of sentiment analyses from Alvarez et al. [22] would render even more robust results. Additional joint content and topology analyses would be even more useful: would using Garcia-Herranz et al. [24]'s central users in communities, i.e., incorporate Weng et al.'s methods [23], result in to more precise virality predictor? Would adding content analysis as used in Alvarez et al. [22] further improve precision? If holistic understanding of social phenomena is researchers goal, future efforts should seek to

incorporate not one but numerous methodologies in pursuit of that end.

5. CONCLUSION

The papers considered in this minireview offer several important considerations on the state of Twitter research into social phenomena. What was once the arena of solely political scientists and sociologists, political action and social phenomena have now become research topics for computer scientists and social physicists. New disciplines have much to offer social research, as indicated in the methodology review of our sample papers; yet, these methodologies are often divorced from underlying social theory. Thus far, Twitter studies offer primarily observational—not explanatory—analyses.

What does account for this bias away from social theory? Some possible explanations are readily apparent. Twitter research is new, and computational social science is an emerging field; thus far both have tended to prioritize methodological innovation over incorporation or analysis of preexisting social theories. This tendency has surely been exacerbated by the relatively narrow range of disciplines contributing to the field: despite its name, the field has drawn from computer

scientists, mathematicians, and physicists far more than social scientists. Perhaps interdisciplinary collaboration may present a solution as the field continues to develop; see Beguerisse-Díaz et al. [41] for a recent example. The tendency to disregard social theory also likely has its origins in the structure of technical journals. A high premium on space and their technical audience simply do not permit lengthy discussion of theory.

Greater dialog between theory and methods, as well as a holistic use of all available methodologies, is needed for data science to truly offer insight into our social world, both on Twitter and off it.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

For providing useful feedback on the original manuscript we thank Mariano Beguerisse-Díaz and Peter Grindrod.

REFERENCES

- Weller K, Bruns A, Burgess J, Mahrt M, Puschmann C. *Twitter and Society*. New York, NY: Peter Lang (2013).
- Margets H, John P, Hale S, Yasseri T. *Political Turbulence: How Social Media Shape Collective Action*. Princeton, NJ: Princeton University Press (2015).
- Ruths D, Pfeffer J. Social media for large studies of behavior. *Science* (2014) **346**:1063–4. doi: 10.1126/science.346.6213.1063
- Steinert-Threlkeld ZC, Mocanu D, Vespignani A, Fowler J. Online social networks and offline protest. *EPJ Data Sci.* (2015) **4**:1. doi: 10.1140/epjds/s13688-015-0056-y
- Kwak H, Lee C, Park H, Moon S. What is Twitter, a social network or a news media? In: *Proceedings of the 19th International Conference on World Wide Web*. New York, NY: ACM (2010). pp. 591–600.
- Huberman BA, Romero DM, Wu F. Social Networks that Matter: Twitter under the Microscope. *First Monday* (2009) **14**:1.
- Gonçalves B, Perra N, Vespignani A. Modeling users' activity on twitter networks: validation of dunbar's number. *PLoS ONE* (2011) **6**:e22656. doi: 10.1371/journal.pone.0022656
- Bakshy E, Hofman JM, Mason WA, Watts DJ. Everyone's an influencer: quantifying influence on twitter. In: *Proceedings of the Fourth ACM International Conference on Web Search and Data Mining*. New York, NY: ACM (2011). pp. 65–74.
- Yan P, Yasseri T. Two roads diverged: a semantic network analysis of guanxi on twitter. arXiv:160505139 (2016).
- Livne A, Simmons MP, Adar E, Adamic LA. The party is over here: structure and content in the 2010 election. *ICWSM* (2011) **11**:17–21. Available online at: <http://www.aaai.org/ocs/index.php/ICWSM/ICWSM11/paper/view/2852>
- Lietz H, Wagner C, Bleier A, Strohmaier M. When politicians talk: assessing online conversational practices of political parties on Twitter. In: *Proceedings of the Eighth International Conference on Weblogs and Social Media*. Ann Arbor, MI: ICWSM (2014).
- Grindrod P, Lee T. Comparison of social structures within cities of very different sizes. *Open Sci.* (2016) **3**:150526. doi: 10.1098/rsos.150526
- Graham M, Hale SA, Gaffney D. Where in the world are you? Geolocation and language identification in Twitter. *Prof Geogr.* (2014) **66**:568–78. doi: 10.1080/00330124.2014.907699
- Jungherr A. Twitter in politics: a comprehensive literature review. SSRN 2402443. (2014). doi: 10.2139/ssrn.2402443
- Gayo-Avello D. "I wanted to predict elections with Twitter and all I got was this lousy paper"—a balanced survey on election prediction using Twitter data. arXiv:12046441 (2012).
- Metaxas PT, Mustafaraj E. Social media and the elections. *Science* (2012) **338**:472–3. doi: 10.1126/science.1230456
- Petticrew M, Roberts H. *Systematic reviews in the social sciences: a practical guide*. Hoboken, NJ: John Wiley & Sons (2008).
- DiMaggio P, Hargittai E, Celeste C, Shafer S. From unequal access to differentiated use: a literature review and agenda for research on digital inequality. In: *Social Inequality*. New York, NY: Russell Sage Foundation (2004). pp. 355–400.
- Beguerisse-Díaz M, Garduño-Hernández G, Vangelov B, Yaliraki SN, Barahona M. Interest communities and flow roles in directed networks: the Twitter network of the UK riots. *J R Soc Interface* (2014) **11**:20140940. doi: 10.1098/rsif.2014.0940
- Borge-Holthoefer J, Magdy W, Darwish K, Weber I. Content and network dynamics behind Egyptian political polarization on Twitter. In: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. New York, NY: ACM (2015). pp. 700–11.
- Conover MD, Gonçalves B, Flammini A, Menczer F. Partisan asymmetries in online political activity. *EPJ Data Sci.* (2012) **1**:1–19. doi: 10.1140/epjds6
- Alvarez R, Garcia D, Moreno Y, Schweitzer F. Sentiment cascades in the 15M movement. *EPJ Data Sci.* (2015) **4**:1–13. doi: 10.1140/epjds/s13688-015-0042-4
- Weng L, Menczer F, Ahn YY. Virality prediction and community structure in social networks. *Sci Rep.* (2013) **3**:2522. doi: 10.1038/srep02522
- García-Herranz M, Moro E, Cebrian M, Christakis NA, Fowler JH. Using friends as sensors to detect global-scale contagious outbreaks. *PLoS ONE* (2014) **9**:e92413. doi: 10.1371/journal.pone.0092413
- Coppock A, Guess A, Ternovski J. When treatments are Tweets: a network mobilization experiment over Twitter. *Pol Behav.* (2016) **38**:105–128. doi: 10.1007/s11109-015-9308-6
- González-Bailón S, Borge-Holthoefer J, Rivero A, Moreno Y. The dynamics of protest recruitment through an online network. *Sci Rep.* (2011) **1**:197. doi: 10.1038/srep00197

27. Vidgen B, Yasseri T. P-values: misunderstood and misused. *Front Phys.* (2016) **4**:6. doi: 10.3389/fphy.2016.00006
28. Wasserstein RL, Lazar NA. The ASA's statement on p-values: context, process, and purpose. *Am Stat.* (2016) **70**:129–33. doi: 10.1080/00031305.2016.1154108
29. Morstatter F, Kumar S, Liu H, Maciejewski R. Understanding Twitter data with TweetXplorer. In: *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. KDD '13*. New York, NY: ACM (2013). pp. 1482–5.
30. Morstatter F, Pfeffer J, Liu H, Carley KM. Is the sample good enough? Comparing data from Twitter's streaming API with Twitter's firehose. In: *Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media*. New York, NY: AAAI Press (2013). p. 400.
31. Holme P, Saramäki J. Temporal networks. *Phys Rep.* (2012) **519**:97–125. doi: 10.1016/j.physrep.2012.03.001
32. Holme P. Modern temporal network theory: a colloquium. *Eur Phys J B* (2015) **88**:1–30. doi: 10.1140/epjb/e2015-60657-4
33. Newman M. *Networks: An Introduction*. Oxford: Oxford University Press (2010).
34. Wasserman S, Faust K. *Social Network Analysis: Methods and Applications*. Vol. 8. Cambridge: Cambridge University Press (1994).
35. González-Bailón S, Wang N. The bridges and brokers of global campaigns in the context of social media. In: *SSRN Work Paper* (2013).
36. Rosvall M, Bergstrom CT. Maps of random walks on complex networks reveal community structure. *Proc Natl Acad Sci USA.* (2008) **105**:1118–23. doi: 10.1073/pnas.0706851105
37. Raghavan UN, Albert R, Kumara S. Near linear time algorithm to detect community structures in large-scale networks. *Phys Rev E* (2007) **76**:036106. doi: 10.1103/PhysRevE.76.036106
38. Newman ME. Finding community structure in networks using the eigenvectors of matrices. *Phys Rev E* (2006) **74**:036104. doi: 10.1103/PhysRevE.74.036104
39. Schaub MT, Delvenne JC, Yaliraki SN, Barahona M. Markov dynamics as a zooming lens for multiscale community detection: non clique-like communities and the field-of-view limit. *PLoS ONE* (2012) **7**:e32210. doi: 10.1371/journal.pone.0032210
40. Fortunato S. Community detection in graphs. *Phys Rep.* (2010) **486**:75–174. doi: 10.1016/j.physrep.2009.11.002
41. Beguerisse-Díaz M, McLennan AK, Garduño-Hernández G, Barahona M, Ulijaszek SJ. The 'who' and 'what' of # diabetes on Twitter. arXiv:150805764 (2015).

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Cihon and Yasseri. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read,
for greatest visibility



COLLABORATIVE PEER-REVIEW

Designed to be rigorous
– yet also collaborative,
fair and constructive



FAST PUBLICATION

Average 85 days from
submission to publication
(across all journals)



COPYRIGHT TO AUTHORS

No limit to article
distribution and re-use



TRANSPARENT

Editors and reviewers
acknowledged by name
on published articles



SUPPORT

By our Swiss-based
editorial team



IMPACT METRICS

Advanced metrics
track your article's impact



GLOBAL SPREAD

5'100'000+ monthly
article views
and downloads



LOOP RESEARCH NETWORK

Our network
increases readership
for your article

Frontiers

EPFL Innovation Park, Building I • 1015 Lausanne • Switzerland
Tel +41 21 510 17 00 • Fax +41 21 510 17 01 • info@frontiersin.org
www.frontiersin.org

Find us on

