Innovations are Disproportionately Likely in the Periphery of a Scientific Network

DERYC T. PAINTER, BRYAN C. DANIELS, MANFRED D. LAUBICHLER, Arizona State University

1. INTRODUCTION

Innovation is a popular idea. One might describe it as a universal goal of businesses, universities, and scientists. Economists and business entrepreneurs see it as a means to sustain growth and increase productivity and revenue [Szirmai et al. 2011; Naranjo-Valencia et al. 2018]. Universities push for the creation of innovative research and teaching styles [Amador et al. 2018; Kryukov and Gorin 2016]. Sociologists and animal behaviorists often explore innovation through new behaviors or ideas that propagate throughout a group [Brosnan and Hopper 2014; Ramsey et al. 2007]. Developmental biologists and geneticists view innovation in terms of genetic variants that spread to become common in a population [Wagner 2014; Davidson and Erwin 2010]. All of these representations of innovation share a common theme: novelty, or invention, followed by adoption and propagation. A new idea or behavior is created and then integrated into the surrounding community.

The process of innovation is inherently collective, requiring shared consensus on novel strategies. In a sense, the process of innovation spreading represents one of the simplest and most ancient algorithms for computing optimal solutions to shared problems. Determining the conditions that contribute to successful innovation is therefore fundamental to an understanding of the broader subject of collective computation.

Here, we focus on identifying conditions that lead to successful innovations in modern science. Some have argued that innovation is easier at the fringes of fields where novel connections can come without being overly constrained by the current status quo [da Motta e Albuquerque 2007; Fitjar and Rodríguez-Pose 2011]. We aim to bring quantitative rigor to this hypothesis, asking: Are innovations indeed more likely to arise from areas outside the mainstream, or are well-established lines of research that are more well-tested in fact better able to inject long-lasting conceptual change?

1.1 An Operational Definition of Innovation

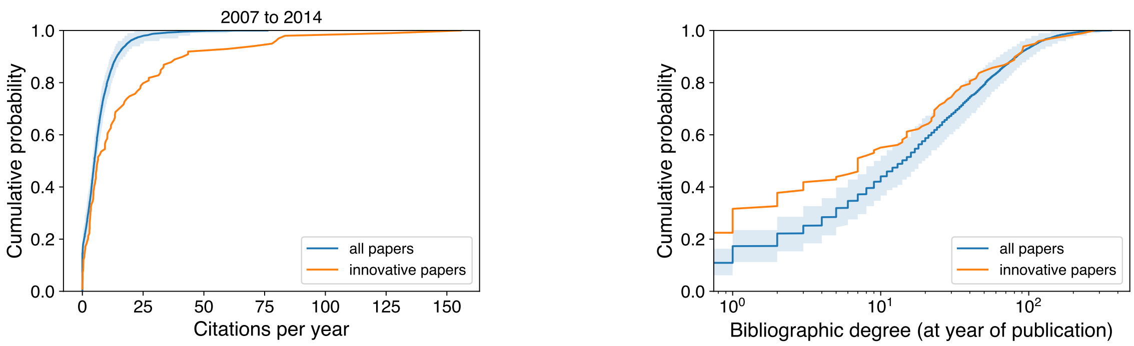
Innovation can be a nebulous term with a variety of interpretations, even within a discipline. Economics, business, and finance often view innovation through a monetary and efficiency lens with a goal of greater sustained profit over time [Szirmai et al. 2011; Naranjo-Valencia et al. 2018]. Behavioral innovation is considered in terms of individual novelty followed by cultural transmission [Brosnan and Hopper 2014; Ramsey et al. 2007]. In a biological context, innovation is often viewed in a genetic and developmental framework [Wagner 2014; Davidson and Erwin 2010]. Novelty and persistence repeat in each of these examples. The theme of novelty and persistence is echoed from Schumpeter's and Brozen's work, "Invention, Innovation, Imitation." [Brozen 1951] Our novelty equates to their idea of invention. Innovation and imitation are considered together in our notion of persistence. We, therefore, base our operational definition of innovation on the joint occurrence of novelty and persistence.

1.2 Evolutionary Medicine

The field of evolutionary medicine is ideal for our study of innovation because it offers a unique opportunity to study innovation within context of the marriage between two distinct scientific fields, evolutionary biology and human health and disease. Evolutionary medicine began with the article ”Dawn of Darwinian Medicine” [Williams and Nesse 1991]. These ideas expanded into an endeavor to better educate clinicians with the general principles of evolution as they relate to human health and disease [Nesse et al. 2009]. The ideological framework of evolutionary medicine is novel and persists today as a collective innovation of sorts built on the individual innovations of its practitioners. Our study examines the language of evolutionary medicine, written by those who self-identify as interested in evolutionary medicine, to computationally identify innovations and their origins across all publications in the field.

1.3 Using Bibliographic Coupling Networks to Characterize Publications

We first characterize conceptual connections among publications in the corpus using a bibliographic coupling network. Two publications are joined by an edge in this network when there is at least one other reference that they both cite. The structure of this network characterizes the relative popularity and connectivity of various concepts: manuscripts with a large degree use concepts that are more closely related to previous work, with clusters of highly-connected manuscripts citing the same popular references [Kessler 1963; Yan and Ding 2012]. We expected concepts to be used heterogeneously, with some popular core concepts that accumulate large clusters of manuscripts, and some peripheral concepts that cite references that are dissimilar to other manuscripts in the corpus. We characterize the degree to which the conceptual network displays a distinct core and periphery using the rich club coefficient [Colizza et al. 2006; Opsahl et al. 2008].



(a) Publications identified as innovative accumulate citations at a much faster rate. Plotted is the cumulative probability function for citations per year attained by Evolutionary Medicine papers published between 2007 and 2014 (measuring the proportion of papers with the given citation rate or lower).

(b) Innovative papers are disproportionately likely in the periphery. The distribution of bibliographic degree at the time of publication (displayed as the cumulative probability of having degree less than or equal to the given value).

Fig. 1: In (a), the 100 papers with largest innovation index I (orange) accumulated an average of 17.2 citations per year, while all papers (blue) accumulated an average of 6.4 citations per year. The 90% confidence intervals for randomly chosen subsets of 100 papers are shown as a blue shaded region. The distributions are highly significantly distinct (KS-test statistic 0.23, $p < 10^{-4}$). In (b), Compared to all publications in the given years (blue), the 100 most innovative papers (orange) are more likely to have small bibliographic degree, indicating they are more likely to lie outside the rich club. As suggested by comparing to randomly chosen sets of 100 papers (90% confidence intervals shown as blue shaded region), the difference between the distributions is statistically significant (two-sample KS-test statistic 0.158, $p = 0.013$).

2. RESULTS

In this paper, we take a step in this direction by defining two quantitative aggregate measures: a measure of innovation and a characterization of the core-periphery structure of concepts within the studied field. With these quantitative measures, we are then able to rigorously test a hypothesis about collective properties that lead to increased innovation.

Our main finding in this work is that innovative publications occur significantly more often than expected in the periphery of the bibliographic coupling network. We conjecture that this may occur due to the flexibility afforded to these publications through their lack of strong connections to other publications. In the core, documents exhibit a high degree of sameness to other documents. The difficulty of producing persistent novelty may be greater when a publication is based on the same concepts used by many others in the field. Conversely, the publications at the periphery are less constrained by the high level of sameness in the core. It is worth noting that innovations do occur in the core, but not at a significantly different rate than would be expected based on the fraction of papers located there.

2.1 Core-periphery Function in Collective Computation

Having found that innovations are over-represented in the periphery of a scientific field, we may speculate about functional advantages of the core-periphery structure that is common to a number of scientific networks. While maintaining a thriving core may be crucial for developing existing ideas, allowing for a diverse periphery could be equally important to avoid stasis and to promote necessary adaptation. It is likely that, as previous work on bibliographic coupling networks suggests [Ferreira 2018; Nettle and Frankenhuis 2019], the paradigm exemplified by a scientific field is, in a sense, defined by the stability of the rich-clubs, and we are more likely to encounter a Kuhn-like scientific revolution [Kuhn 1962] originating in the more volatile, interdisciplinary periphery.

More broadly, trade-offs between robustness and adaptability occur across a range of biosocial systems. For instance, living systems may tune their distance from a collective instability in order to be either more predictable and robust (further from the transition) or sensitive and adaptable (closer to the transition) [Daniels et al. 2017]. The core-periphery structure may be viewed as a strategy for specialization of different parts of the network to different levels of flexibility.

In neuroscience, a clear rich-club structure has been found within brain networks [Bassett et al. 2013]. There is some evidence that this structure could allow for both robust stereotyped learned behavior (controlled by the core) and higher variance behavior needed for faster adaptation (occurring in the periphery). In particular, highly connected core areas are found to be slower to change during learning compared to peripheral areas [Bassett et al. 2013], and a detailed computational model of the macaque brain displays stable dynamics in the core simultaneous with unstable dynamics in the periphery [Gollo et al. 2015]. Looking forward, we may be inspired to look for a similar dynamical pattern in scientific networks in which the rich-club is able to retain robust knowledge by virtue of its slowness to change.

REFERENCES

Soleidy Rivero Amador, Maidelyn Díaz Pérez, María José López-Huertas, and Reinaldo Javier Rodríguez Font. 2018. Indicator system for managing science, technology and innovation in universities. *Scientometrics* 115, 3 (2018), 1575–1587. DOI:http://dx.doi.org/10.1007/s11192-018-2721-y

Danielle S. Bassett, Nicholas F. Wymbs, M. Puck Rombach, Mason A. Porter, Peter J. Mucha, and Scott T. Grafton. 2013. Task-Based Core-Periphery Organization of Human Brain Dynamics. *PLoS Computational Biology* 9, 9 (2013), 1–16. DOI:http://dx.doi.org/10.1371/journal.pcbi.1003171

Sarah F Brosnan and Lydia M Hopper. 2014. Psychological limits on animal innovation. *Animal Behaviour* 92 (2014), 325–332. DOI:http://dx.doi.org/10.1016/j.anbehav.2014.02.026

Yale Brozen. 1951. Invention, innovation, and imitation. *The American Economic Review* 41, 2 (1951), 239–257.

Vittoria Colizza, Alessandro Flammini, M. Angeles Serrano, and Alessandro Vespignani. 2006. Detecting rich-club ordering in complex networks. *Nature Physics* 2, February (2006), 110. DOI:http://dx.doi.org/10.1038/nphys209

Eduardo da Motta e Albuquerque. 2007. Inadequacy of technology and innovation systems at the periphery. *Cambridge Journal of Economics* 31, 5 (2007), 669–690. DOI:http://dx.doi.org/10.1093/cje/bel045

Bryan C. Daniels, David C. Krakauer, and Jessica C. Flack. 2017. Control of finite critical behaviour in a small-scale social system. *Nature Communications* 8 (2017), 14301. DOI:http://dx.doi.org/10.1038/ncomms14301

Eric H. Davidson and Douglas H. Erwin. 2010. Evolutionary innovation and stability in animal gene networks. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 314 B, 3 (2010), 182–186. DOI:http://dx.doi.org/10.1002/jez.b.21329

Fernando AF Ferreira. 2018. Mapping the field of arts-based management: Bibliographic coupling and co-citation analyses. *Journal of Business Research* 85 (2018), 348–357.

Rune Dahl Fitjar and AndréS Rodríguez-Pose. 2011. Innovating in the periphery: Firms, values and innovation in Southwest Norway. *European Planning Studies* 19, 4 (2011), 555–574. DOI:http://dx.doi.org/10.1080/09654313.2011.548467

Leonardo L. Gollo, Andrew Zalesky, R. Matthew Hutchison, Martijn van den Heuvel, and Michael Breakspear. 2015. Dwelling quietly in the rich club: Brain network determinants of slow cortical fluctuations. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370, 1668 (2015), 20140165. DOI:http://dx.doi.org/10.1098/rstb.2014.0165

M. M. Kessler. 1963. Bibliographic coupling between scientific papers. *American Documentation* 14, 1 (1963), 10–25. DOI:http://dx.doi.org/10.1002/asi.5090140103

V Kryukov and A Gorin. 2016. Digital Technologies As Education Innovation At Universities. *Journal of Internet Banking and* 21, 3 (2016), 1–18. http://search.proquest.com/openview/7f1bd6e899ab69ef9b6f96345ff2af55/1?pq-origsite=gscholar

Thomas S Kuhn. 1962. The Structure of Scientific Revolutions. *Philosophy and Rhetoric of Science* 73, 3 (1962), 65.

J. C. Naranjo-Valencia, G. Calderon-Hernandez, D. Jimenez-Jimenez, and R. Sanz-Valle. 2018. Entrepreneurship and Innovation: Evidence in Colombian SMEs. In *Handbook of Research on Intrapreneurship and Organizational Sustainability in SMEs*. IGI Global, PO Box 12205, Arlington, VA 22219-2205, 294–316.

Randolph M Nesse, Carl T Bergstrom, Peter T Ellison, Jeffrey S Flier, Peter D Gluckman, Diddahally R Govindaraju, Dietrich Niethammer, Gilbert S Omenn, Robert L Perlman, Mark D Schwartz, Mark G Thomas, Stephen C Stearns, and David Valle. 2009. Making evolutionary biology a basic science for medicine. *PNAS* 107 (2009), 1–8. http://scholar.google.com/scholar?start=40

Daniel Nettle and Willem E Frankenhuis. 2019. The evolution of life-history theory: a bibliometric analysis of an interdisciplinary research area. *Proceedings of the Royal Society B* 286, 1899 (2019), 20190040.

Tore Opsahl, Vittoria Colizza, Pietro Panzarasa, and José J. Ramasco. 2008. Prominence and control: The weighted rich-club effect. *Physical Review Letters* 101, 16 (2008), 1–4. DOI:http://dx.doi.org/10.1103/PhysRevLett.101.168702

Grant Ramsey, Meredith L Bastian, and Carel Van Schaik. 2007. Animal innovation de ned and operationalized. *Innovation* 30, 4 (2007), 393 –437. DOI:http://dx.doi.org/10.1017/S0140525X07002373

Adam Szirmai, Wim Naude, and Micheline Goedhuys. 2011. *Entrepreneurship, innovation, and economic development*. Oxford University Press, 198 Madison Avenue, New York, NY 10016 USA. DOI:http://dx.doi.org/10.1093/acprof

Günter P. Wagner. 2014. *Homology, Gene, and Evolutionary Innovation*. Princeton University Press, 41 William Street, Princeton, New Jersey.

George C Williams and Randolph M Nesse. 1991. The Dawn of Darwinian Medicine. *Quarterly Review of Biology* 66, 1 (1991), 1–22.

Erjia Yan and Ying Ding. 2012. Scholarly network similarities: How bibliographic coupling networks, citation networks, cocitation networks, topical networks, coauthorship networks, and coword networks relate to each other. *Journal of the American Society for Information Science and Technology* 63, 7 (2012), 1313–1326. DOI:http://dx.doi.org/10.1002/asi.22680