Eigenfactor: ranking and mapping scientific knowledge

Jevin D. West
University of Washington
jevinw@u.washington.edu
The Eigenfactor Algorithm

**Q:** How can we better evaluate the scholarly literature?

The Eigenfactor Project

**Q:** How does network structure affect function?
Citation Networks, Scaling Up

Network Effects, Centrality

Eigenfactor, Article Influence

Mapping, Microsoft Academic Search
Citation Networks, Scaling Up
Coevolutionary cycling of host sociality and pathogen virulence in contact networks

Federico Prado, Alyssa Sheil, Devin W. West, Benjamin Kerr*

Institute of Biomedical Sciences, University of Washington, Box 357220, Seattle, WA 98195, USA

ARTICLE INFO
Abstract

In many animal species, groups forming with significant advantages to the individual comprising the group, including protection from parasites, learning and foraging efficiency, enhanced information exchange, reduced energy expenditure in movement and thermoregulation, and improved access to mates and helpers for infanticide (Alexander, 1974; Beauchamp, 2004; Caraco and Keesling, 2001; Higdon and Sherman, 1976; Kacser and Burns, 2002; Latour, 1976; Lee 1984; Pulliam, 2000; Snowdon, 1975). However, group living comes with costs for group members as well as benefits. Group members may experience increased competition for food, mating, and mates among members of a group may lead to inbreeding depression, group-level fluctuations, and increased likelihood of related infections and reduced transmission in non-kin. In this study, we examine the effects of group size and group structure on the transmission of parasites within and between groups, and the impact of parasite transmission on the evolution of group membership. We find that group size and group structure can affect the transmission of parasites, and the impact of parasite transmission on the evolution of group membership. We also find that parasite transmission can affect the evolution of group membership, and the impact of parasite transmission on the evolution of group membership. We conclude that parasite transmission can affect the evolution of group membership, and the impact of parasite transmission on the evolution of group membership.
Citations

Evolution of Virulence: A Unified Framework for Infection and Superinfection

J. Mosquera* and Frederick R. Adler†

*Group of Nonlinear Physics, Faculty of Physics, University of Santiago de Compostela, 15706 Santiago de Compostela, Spain and University Department of Mathematics and Department of Biology, Salt Lake City, UT 84112, U.S.A.

(Received on 28 November 1997, Accepted in revised form on 21 July 1998)

Abstract: We develop a model of infection and superinfection based on the idea that hosts infected with one strain of a parasite may be more susceptible to infection by a second strain. We analyze the model and show that, for some parameter values, the model exhibits a Hopf bifurcation, leading to the emergence of oscillations in the number of infected hosts. We also show that the model is capable of exhibiting a variety of other dynamical behaviors, including, for example, the emergence of periodic oscillations in the number of infected hosts. We conclude that the model is capable of exhibiting a wide range of dynamical behaviors, and that it may be useful for understanding the evolution of virulence in parasitic infections.
Scholarship is all about the flow of ideas
Networks are all about the flow of information
Citations form a vast network

de Solla Price, Science (1965)
Citations among Authors

Eugene Garfield

Derek de Solla Price
What happens when you scale up?

12,000 Journals
150,000,000 Citations
High School Dating Network

Image Courtesy of Mark Newman
High School Dating Network

System Scale: Network Topology

Q: Which network would MONO prefer?

Individual Scale: Connection Source

Q: Which girl/boy should I date if I don’t want MONO?
Who is the most ‘influential’ high school student?
Degree Centrality

Eigenvector Centra
Degree Centrality

Number of students

Links

1 2 3 4 5 6 7 8 9 10 11 12 13 14
Accounting for the Network
The Network Matters
This network property of the scholarly literature was largely ignored over the first century of scholarly evaluation.

How can we extract this information in order to better measure information flow?
Impact factor

Cites in 2010 to articles in 2009 or 2008

Number of articles in 2009 and 2008

Garfield, Science (1955)
Impact Factor

2010

nature, Cell, Science, PNAS, Ecology

3 citations

2008 - 2009

Plant Physiology

2 articles

\[ IF_{2010} = \frac{3}{2} \]
Impact Factor

Eigenfactor
Eigenfactor algorithm

\[ P = \alpha H + (1 - \alpha) \text{a.e}^T \]

Matrix representing the random walk over citations

Probability of not teleporting

Cross-citation Matrix dictating the structure of the citation network

Probability of teleporting to completely new journal weighted by the number of articles in that journal

\[ EF = 100 \frac{H}{_i[H]} \]

Leading eigenvector of the random walk matrix \( P \).

Normalization

Bergstrom (2007); West et al (2010)
Which node is the most central?

25 nodes and 42 weighted, directed links
Deterministic: iterative voting
Stochastic: random walker

Applet: Daniel Edler

<table>
<thead>
<tr>
<th>Rank</th>
<th>Journal</th>
<th>Eigenfactor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Nature</td>
<td>1.76</td>
</tr>
<tr>
<td>2</td>
<td>PNAS</td>
<td>1.70</td>
</tr>
<tr>
<td>3</td>
<td>Science</td>
<td>1.58</td>
</tr>
<tr>
<td>4</td>
<td>Journal of Biological Chemistry</td>
<td>1.33</td>
</tr>
<tr>
<td>5</td>
<td>Physical Review Letters</td>
<td>1.28</td>
</tr>
<tr>
<td>6</td>
<td>Journal of the American Chemical Society</td>
<td>0.95</td>
</tr>
<tr>
<td>7</td>
<td>Physical Review B</td>
<td>0.77</td>
</tr>
<tr>
<td>8</td>
<td>Applied Physics Letters</td>
<td>0.72</td>
</tr>
<tr>
<td>9</td>
<td>New England Journal of Medicine</td>
<td>0.68</td>
</tr>
<tr>
<td>10</td>
<td>Cell</td>
<td>0.67</td>
</tr>
</tbody>
</table>

Total number of journals ranked = 8,232

<table>
<thead>
<tr>
<th>Rank</th>
<th>Journal</th>
<th>Eigenfactor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Nature</td>
<td>1.76 %</td>
</tr>
<tr>
<td>2</td>
<td>PNAS</td>
<td>1.70 %</td>
</tr>
<tr>
<td>3</td>
<td>Science</td>
<td>1.58 %</td>
</tr>
<tr>
<td>4</td>
<td>Journal of Biological Chemistry</td>
<td>1.33 %</td>
</tr>
<tr>
<td>5</td>
<td>Physical Review Letters</td>
<td>1.28 %</td>
</tr>
<tr>
<td>6</td>
<td>Journal of the American Chemical Society</td>
<td>0.95 %</td>
</tr>
<tr>
<td>7</td>
<td>Physical Review B</td>
<td>0.77 %</td>
</tr>
<tr>
<td>8</td>
<td>Applied Physics Letters</td>
<td>0.72 %</td>
</tr>
<tr>
<td>9</td>
<td>New England Journal of Medicine</td>
<td>0.68 %</td>
</tr>
<tr>
<td>10</td>
<td>Cell</td>
<td>0.67 %</td>
</tr>
</tbody>
</table>

Total number of journals ranked = 8,232
Article Influence = \frac{\text{Eigenfact}}{\# \text{ articles}}

<table>
<thead>
<tr>
<th>Rank</th>
<th>Journal</th>
<th>Article Influence</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Review of Modern Physics</td>
<td>24.9</td>
</tr>
<tr>
<td>2</td>
<td>Annual Review of Immunology</td>
<td>24.7</td>
</tr>
<tr>
<td>3</td>
<td>Annual Review of Biochemistry</td>
<td>20.9</td>
</tr>
<tr>
<td>4</td>
<td>Nature Reviews Molecular Cell Biology</td>
<td>20.0</td>
</tr>
<tr>
<td>5</td>
<td>Annual Review of Neuroscience</td>
<td>18.9</td>
</tr>
<tr>
<td>6</td>
<td>Cell</td>
<td>18.9</td>
</tr>
<tr>
<td>7</td>
<td>New England Journal of Medicine</td>
<td>18.8</td>
</tr>
<tr>
<td>8</td>
<td>CA – A Cancer Journal for Clinicians</td>
<td>17.5</td>
</tr>
<tr>
<td>9</td>
<td>Nature</td>
<td>17.3</td>
</tr>
<tr>
<td>10</td>
<td>Science</td>
<td>16.3</td>
</tr>
</tbody>
</table>

Total number of journals ranked = 8,232
start a search

search by
journal name

Eigenfactor™ subject category
year

Use advanced search to search by Thomson JCR subject categories, publisher, and other fields.

### Journal Summary List

#### Journals from: subject categories EVOLUTIONARY BIOLOGY

**Sorted by:** Article Influence (TM) Score

#### Journals 1 - 20 (of 39)

**Ranking is based on your journal and sort selections.**

<table>
<thead>
<tr>
<th>Mark</th>
<th>Rank</th>
<th>Abbreviated Journal Title (linked to journal information)</th>
<th>ISSN</th>
<th>Total Cites</th>
<th>Impact Factor</th>
<th>5-Year Impact Factor</th>
<th>Immediacy Index</th>
<th>Articles</th>
<th>Cited Half-life</th>
<th>Eigenfactor Score</th>
<th>Article Influence Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>ANNU REV ECOL EVOL S</td>
<td>1543-592X</td>
<td>10653</td>
<td>10.161</td>
<td>17.176</td>
<td>0.133</td>
<td>30</td>
<td>&gt; 10.0</td>
<td>0.02146</td>
<td>8.322</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>TRENDS ECOL EVOL</td>
<td>0169-5347</td>
<td>16830</td>
<td>11.904</td>
<td>17.188</td>
<td>1.913</td>
<td>92</td>
<td>7.4</td>
<td>0.06469</td>
<td>7.846</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>SYST BIOL</td>
<td>1063-5157</td>
<td>7636</td>
<td>7.833</td>
<td>12.743</td>
<td>0.946</td>
<td>74</td>
<td>5.9</td>
<td>0.03104</td>
<td>4.668</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>MOL BIOL EVOL</td>
<td>0737-4038</td>
<td>20661</td>
<td>7.280</td>
<td>6.846</td>
<td>1.259</td>
<td>263</td>
<td>7.0</td>
<td>0.08141</td>
<td>2.951</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>AM NAT</td>
<td>0003-0147</td>
<td>21622</td>
<td>4.670</td>
<td>5.508</td>
<td>0.784</td>
<td>190</td>
<td>&gt; 10.0</td>
<td>0.04815</td>
<td>2.586</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>EVOLUTION</td>
<td>0014-3820</td>
<td>25826</td>
<td>4.737</td>
<td>5.427</td>
<td>0.798</td>
<td>247</td>
<td>&gt; 10.0</td>
<td>0.06053</td>
<td>2.197</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>BMC EVOL BIOL</td>
<td>1471-2148</td>
<td>2562</td>
<td>4.050</td>
<td>4.373</td>
<td>0.508</td>
<td>331</td>
<td>2.2</td>
<td>0.02196</td>
<td>1.888</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>MOL ECOL</td>
<td>0962-1083</td>
<td>19960</td>
<td>5.325</td>
<td>5.966</td>
<td>1.506</td>
<td>403</td>
<td>5.5</td>
<td>0.06926</td>
<td>1.811</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>PALEOBIOLOGY</td>
<td>0094-8373</td>
<td>2761</td>
<td>2.800</td>
<td>3.063</td>
<td>0.688</td>
<td>32</td>
<td>&gt; 10.0</td>
<td>0.00728</td>
<td>1.739</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>CLADISTICS</td>
<td>0748-3007</td>
<td>2966</td>
<td>3.515</td>
<td>4.881</td>
<td>0.647</td>
<td>51</td>
<td>&gt; 10.0</td>
<td>0.00576</td>
<td>1.625</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>J EVOLUTION BIOL</td>
<td>1010-061X</td>
<td>5691</td>
<td>3.471</td>
<td>3.757</td>
<td>0.767</td>
<td>180</td>
<td>5.0</td>
<td>0.02980</td>
<td>1.556</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>EVOL DEV</td>
<td>1520-541X</td>
<td>1424</td>
<td>3.637</td>
<td>3.623</td>
<td>0.634</td>
<td>71</td>
<td>4.5</td>
<td>0.00829</td>
<td>1.513</td>
</tr>
</tbody>
</table>
Mapping
good maps simplify and highlight relevant structures
The Map Equation

\[ L(M) = q \bowtie H(Q) + \sum_{i=1}^{m} p_i \bowtie H(P^i) \]

The map equation tells us the description length for a particular modular structure.
Rovall & Bergstrom (2008)
Hierarchical Maps

Rolvall & Bergstrom (2011)
Rovall & Bergstrom (2010)
Future Directions
How can we better evaluate the scholarly literature?
How can we better navigate the scholarly literature?
Eigenfactor and Microsoft Academic Research
By uncovering the hierarchical structure of scholarly citation, we can identify key papers pertaining to any search query. For a reader new to the field we can find the classic and foundational papers; for an expert we can find the latest innovations.

From patterns of scholarly citation, we use Rosvall and Bergstrom’s map equation to chart the topography of science and the relations among fields and subfields. [journal map] [paper map]

By integrating a hierarchical clustering of citation networks with semantic analysis, we develop a scalable map of scientific fields and the key research terms and topics therein.

Scientific influence is often quantified using simple citation counts, but the structure of a citation network provides far more information than can be revealed by these simple counts. This is principle behind the Eigenfactor metrics; we can better rank the importance of scientific journals or papers by viewing them in the context of the full citation network.
Three things to remember...

1. The scholarly literature forms a vast network

2. There is a wealth of information in the _structure_ of networks

3. Eigenfactor is a _network_ metric that ranks scholarly journals