Links in science : linking network and bibliometric analyses in the study of research performance

Calero Medina, C.M.

Author: Calero Medina, Clara
Title: Links in science : linking network and bibliometric analyses in the study of research performance
Date: 2012-12-11

7 Seed journal citation network maps: A method based on network theory

C.M. Calero Medina , T.N. van Leeuwen (published in the Journal of the American Society for Information Science and Technology, 63(6), 1226-1234, 2012)

7.1 Introduction

Structuring science is about identifying fields, subfields, and research themes and relating them to each other. It is necessary because the traditional science classification system is imperfect, especially for highly multidisciplinary environments, and because it helps to assess performance within its proper context.

In recent years there has been an enormous development in the field of information science in applying different techniques to visualize and analyze the growth of specialties, the structure of scientific communities, and the flow of scientific information (Scharnhorst & Thelwall, 2005).

In fact, as Van Raan (2008) pointed out, science can be considered as an ecosystem comprising species (e.g., fields) whose interdependency can be mapped. The mapping of scientific documents is done in many different ways, depending on the techniques and the purpose on the analysis in which the map is going to be used.

Börner, Chen, and Boyack (2003) reviewed the literature of in bibliometric mapping based on the unit of analysis. The unit of analysis can be documents, relevant terms or words, authors, and journals. Documents are used to visualize and map a knowledge domain with different purposes like analysis of the domain (e.g., Small, 1999) or assessing research performance in a policy context (e.g., Noyons, Moed, & Luwel, 1999). The coword maps are used to unravel the cognitive structure of a field (e.g., Calero, Buter, Cabello, & Noyons, 2006). Authors-based maps are used to infer the intellectual structure of a field (e.g., Chen, 1999). Finally a map of journals can be used to obtain a macro view of science (e.g., Bassecoulard & Zitt, 1999) or to show fine distinctions within a discipline (e.g., Leydesdorff, 1994).

More related to the objective and approach presented in this study is the work done by Leydesdorff and colleagues. In recent years, they have presented a methodology to visualize the citation-impact environment of a given journal (Bornmann, Leydesdorff, & Marx, 2007; Leydesdorff, 2007; Zhou & Leydesdorff, 2007). Their approach makes a distinction between the citing and cited dimensions as two different perspectives on a journal's position (Leydesdorff, 2007). Based on the previous work of He and Pao (1986) and Leydesdorff (1986), the relevant environment for each seed journal (journal under study) is determined by including all journals that cite or are cited by the seed journal to the extent of 1% of its citation rate in the respective dimension. These authors chose the cosine between two vectors (Salton & McGill, 1983) as the similarity measure between the distributions for the various journals included on the citation environment. Visualization is based on social network analysis techniques.

But in understanding and going a step further in the development of visual maps, we can apply network theory. Scientific documents are interconnected trough citations and coauthorships. The seminal work of Derek de Solla Price (1965) showed the structure of science as a network of interconnected publications. We can explore network structures with the help of complex network theory. In recent years researchers, mainly physicists, have started to use the principles of statistical mechanics to analyze large net-worked structures, including science itself (Albert & Barabási, 2002; Dorogovtsev & Mendes, 2002; Newman, Barabási, & Watts, 2006); thus, network techniques are gradually being applied more intensively in bibliometric analysis. Mapping-interrelated entities enables the study of the topology of complex networks. In science such entities are publications, citations (Menczer, 2004; Van Raan, 2005), journals (Bergstrom, West, & Wiseman, 2008), institutes, and authors (Börner, Maru, & Goldstone, 2004).

7.2 Objectives

Traditional quantitative bibliometric indicators are the standard choice nowadays for assessing the research output of a researcher, research group, and research organization (Moed, de Bruin, & van Leeuwen, 1995). But we have to consider that researchers, historians of science, journal editors, librarians, and science managers are also interested in "larger scale questions" that require assessing hundreds or thousands of research papers by a similar number of authors.

From the perspective of bibliometrics and, particularly, journal performance, our goal with journal-citation network analysis is to be able to provide a quick overview of relevant journals related to a journal under study ("seed journal"), in terms of citations given and received. First, it needs to be established what these journals are, how important they might be, and which position they occupy in the network.

As a starting point, we focus on a specific journal, which is considered the seed journal. This seed journal will have citation links with other journals, both given and received (citing to and cited by). When we have a set of journals, we are able to determine the connections between them based on the citations they give and receive. After that we will extract the most prominent journals using a centrality algorithm developed by

Kleinberg (1999) to separate web pages into authorities and hubs. In our analysis, an important authority journal (with high authority centrality weight) is an important source of scientific knowledge in a given set of journals. An important hub journal (with a high hub centrality weight) is an important source of information to look for the most important authority journals. A journal can have both a high hub centrality weight and a high authority centrality weight at the same time: an important source of scientific knowledge and important source of information to look for the most important authority journals.

Finally, we create a network map that comprises the most important hubs and authorities journals related to the seed journal. In just one network map, we will get the relevant citation environment of a specific seed journal. This approach is new because it considers at the same time the citing and cited dimension of a given journal and uses an algorithm developed in complex network theory to detect the prominent journals. These journal citation network graphs are useful for the various stakeholders in and around the science system, as they provide information on the level of journal connections, unlike the more traditional structures these people are familiar with, such as the Journal Subject Categories, the classification system applied in the products of Thomson Reuters (Journal Citation Reports, Web of Science [WoS], etc.). These network graphs clearly show the closest relations journals can have, based on citation relations, suggesting influence relations between journals in such a way that traditional field boundaries are transcended.

7.3 Methods

In this section we present the data and methods used for this study.

Data

In this study, we start from our CWTS in-house database derived from WoS versions of the Science Citation Index and associated citation indices: the Science Citation Index (SCI), the Social Sciences Citation Index (SSCI), and the Arts & Humanities Citation Index (A&HCI). We took all source publications ("articles," "letters," and "reviews") for 2006. The large dataset we created comprises journal-to-journal-citation relations, and it is extracted from the WoS in such a way that all the citing relations of 2006 publications are aggregated to journal level. This means that we grouped the references of (i.e., citations from) the publications in the 2006 source journals to other source publications in the

WoS for the period 1981–2006. So, a citing relation between two journals means a reference (citation given) in 2006 to any other "earlier" publication in the WoS covering the 1981–2006 period. These citations were given by 8,524 journals, citing 8,511 journals covered in the 1981–2006 period. On these data available for the journal citation network analysis, we performed two limiting actions. First we limited the period of citation relations available in the analysis from 1981–2006 to 1997–2006. We monitor then a specific journal (seed journal) and its relations to other journals during a 10-year period (1997–2006). A second limiting action was the creation of symmetry in the dataset, by taking out those journals that are only cited, and not citing journals.

The first step, limiting the period of analysis available in the journal citation network analysis has the following consequences. Initially, this dataset comprised 2,289,383 journal-to-journal relations based on a total of 21,648,745 citations. Limiting the data from the full period to the 10-year period resulted in a dataset of 1,916,714 journal-to-journal relations, and in total 15,528,891 citations. In general, most citations are accounted for within a 10-year window, but it is important to mention that especially for journals in the social sciences and humanities, this limitation is cutting off a larger share of their total number of citations as compared with journals in the natural, life, and technical sciences (Nederhof, 2006). So, the limitation of the period to the 10 most recent years from the year 2006 perspective leads to a loss of 25% of the citations, related to the 1981–1996 period.

Because we want to work with the same journals on the cited range as we work with on the citing range, we limited ourselves to the citing perspective (as this covers all source publications from 2006). In practice, this means that if a journal appeared as a "cited journal" only, we eliminated it from the set. If we then take the next limiting step, the creation of a square matrix of journal-to-journal citation relationships over the citing and cited dimension, for the period 1997–2006, we then started with 8,507 possible journals from which 16 of them appear only as cited by. We removed them from the dataset. This reduction scarcely influences the analysis. This means that at the end what we have is a dataset based on a square matrix (8,491x8,491) of journal-to-journal citation relationships over the citing and cited dimension of 8,491 journals.

Overall, we have asymmetry in the datasets, which comprises two different aspects: first, the asymmetry in the time perspective: the citing year is 2006, cited years are 1997–2006; the second asymmetry comprises the citing-cited relations itself, because a journal can be cited by another, but this does not have to be the other way around. This creates an asymmetrical matrix, in which the upper part is filled differently than the lower part.

"Seed Journal" citation network

For each seed journal we create a matrix that comprises the journal itself, the journals receiving citations from and giving citations to the seed journal, and all citation connections between these other journals. This is what we called the Seed Journal Citation Network. In terms of network theory, the Seed Journal Citation Network is an "ego network." An ego network comprises a focal node ("ego") and the nodes to which the ego node is directly connected to (these are called "alters") plus the ties, if any, among the alters (Hanneman & Riddle, 2005). In the Seed Journal Citation Network, the nodes are individual journals and the edges are values according to how frequently articles published in one journal (positioned in a row) cite articles published in another journal (positioned in a column). The citations are directional (edges) because a citation from journal B to journal A differs from a citation from A to B (this is the asymmetry of the matrix mentioned above). But there are limitations of using only absolute numbers of citations. In particular, they do not reflect the fact that each number on a cell of the seed journal citation matrix depends on the total number of citations given to and received by the two journals. Thus, we developed an index to measure the relationship between two pairs of journals that controls this bias.

Journal Relationship Measure, L index.

Journal citation rates have been used since the seventies to classify journals and delineate specialty fields (Narin, Carpenter, & Berlt, 1972; Narin & Carpenter, 1973; Leydesdorff, 1994; Narin, Hamilton, & Olivasto, 2000; Pudovkin, 1993; Pudovkin & Fuseler, 1995; Pudovkin & Garfield, 2002). However, none of theses approaches consider at the same time the citing and the cited dimension. The approach we present below takes both into account.

The L index reflects the two dimensions of the matrix "citing" and "cited" and considers the global position of the journals in the Web of Science in terms of total citations given and received. Let C_{BA} be the total number of citations given by journal B to Journal A (or what is the same, the total number of citations received by Journal A from Journal B). Let 'TcitingB' be the total number of citations given by Journal B (in the Web of Science) in 2006 and let be 'TcitedA' the total number of citations received by Journal A between 1997-2006. The L index is

$$L_{BA} = \frac{C_{BA}}{\sqrt{TcitingB*TcitedA}}$$

The L index weights the citations given and received. The citations given from one journal to another are weighted by the total number of citations given by that journal and the total number of citations received by that journal. The L index takes values in the interval [0,1]. It is undefined if the total number of citations given by Journal B or received by Journal A is 0. When the number of citations given by Journal B to Journal A is zero, then the measure is 0. The Index reaches its maximum value of 1, when C_{BA}=TcitingB=TcitedA.

Hubs and Authorities

In network theory, a specific research theme focuses on the identification of important nodes in networks. Garfield's impact factor (Garfield, 1972) is a ranking measure based on counting of in-degrees nodes in a journal citation network. Later, Pinski and Narin (1976) and Geller (1978) developed an algorithm that considered not only the number of citations from one journal to the other but also the prestige of the citing journal. Journals that receive many citations from prestigious journals are considered highly prestigious themselves. By iteratively passing prestige from one journal to the other, a stable solution is reached that reflects the relative prestige of journals (Bollen, Rodriguez, & van de Sompel, 2006). This way of measuring prestige is behind the PageRank algorithms to evaluate the status of web pages, first developed by the founders of the Google Search Engine, Brin and Page (Brin & Page, 1998; Page, Brin, Motwani, & Winograd, 1998). The PageRank is calculated through an iterative algorithm that propagates prestige values from one web page to another and converges to a solution (Pillai, Suel, & Cha, 2005).

At the same time Brin and Page created their Google Search Engine, Kleinberg (1999) constructed an algorithm to increase the effectiveness of Web search engines using the concepts of hubs and authorities. Hubs & Authorities are formal notions of structural prominence of vertices in directed graphs (Brandes & Willhalm, 2002). Following Newman (2010), the centrality algorithm developed by Kleinberg is based on the idea that: "there are really two types of important nodes in a directed network: authorities are nodes that contain useful information on a topic of interest; hubs are nodes that tell us where the best authorities are to be found. (page 179)". An authoritative journal, in our case, is one that is cited by many other journals. This idea can be reinforced by observing that citations from all journals aren't equally valuable – some journals are better hubs (citing journals) for a given journal. The algorithm gives each node in a network an authority centrality weight and a hub centrality weight. For each journal (j) in a seed citation network we computed two weights: hub centrality weight (hj) and authority centrality weight (aj). The weights show the strength of a given journal as an authority and/or a hub. Weights are computed according to the citation network (M) by solving the eigenvector problem of matrices MM^T (hubs) and M^TM (authorities), where M is the seed citation matrix (Kleinberg 1999). Journal x is considered a more important hub than journal y if $h_x > h_y$. Journal x is considered a more important authority than journal y if $a_x > a_y$. A node with high authority centrality weight is that it is pointed to by many other vertices with high hub centrality weight. And the characteristic of a node with high hub weight is that it points to many nodes with high authority centrality weight (Newman, 2010).

Kleinberg showed examples in which the algorithm could help filter out irrelevant or poor-quality documents (they would have low authority centrality weights) and to identify high-quality documents (they would have high authority centrality weights). Kleinberg (1999) argued that the tradition of the peer review process in scientific journals ensures that the highly authoritative journals with a common purpose reference one another extensively. He considered then that a one-level model (like the one developed by Pinski and Narin, 1976 and Geller, 1978), in which authorities directly endorse other authorities, fits very well. As mentioned above we are analyzing the whole citation environment of a journal, citing and cited dimension together. From our perspective, making a classification in hubs and authorities is a very useful tool to understand the role played by a journal in the citation environment of a seed journal. An important authority journal (with high authority centrality weight) is an important source of scientific knowledge in a given set of journals. An important hub journal (with a high hub centrality weight) is an important source of information to look for the most important authority journals. A journal can have both a high hub centrality weight and a high authority centrality weight at the same time: an important source of scientific knowledge and important source of information to look for the most important authority journals. This is the reason why we decided to use Kleinberg's algorithm for identifying the main journals in the seed citation network. Batagelj adapted for the software Pajek[1] the Kleinberg's hubs/authorities algorithm (Batagelj & Mrvar, 2006). The results from the analysis presented in this article are based on Pajek.

[1] Pajek is a program for Windows, for analysis and visualization of large networks. It was developed by Vladimir Batagelj and Andrej Mrvar. Some procedures were contributed also by Matjaž Zaveršnik.

7.4 Results

To show the results of our method, we have chosen four journals: Scientometrics, Physical Review Letters, Journal of Vascular and Interventional Radiology, and Public Health. The first journal, Scientometrics, is concerned with the quantitative features and characteristics of science. Emphasis is placed on investigations in which the development and mechanism of science are studied by statistical mathematical methods. The second journal selected, Physical Review Letters, is one of the world's foremost physics journals, providing rapid publication of short reports of significant basic research in all fields of physics. International in scope, this journal provides its diverse readership with weekly coverage of major advances in physics and cross-disciplinary developments. The third journal, Journal of Vascular and Interventional Radiology, is the official journal of the Society of Interventional Radiology. Radiologists, cardiologists, vascular surgeons, neurosurgeons, and other clinicians who need current and reliable information on every aspect of vascular and interventional radiology use it. Each issue covers the most critical medical, minimally invasive, radiological, pathological, and socioeconomic issues of importance to vascular and interventional radiologists. The last journal selected is Public Health, a journal aiming at all public health practitioners and researchers and those who manage public health services and systems.

As was described in the previous section, first we selected the journals in the citation environment for each of the four journals analyzed. The selection is based on the journals receiving citations from and giving citations to the seed journal. Table 1 shows the number of journals selected for each of the four journals. The differences among the four journals analyzed already show characteristics of each of these journals. Scientometrics has 271 journals in its citation environment, showing that it is a very specialized journal in certain types of analyses and data. On the other side, Physical Review Letters has 979 journals, showing that it is a general journal in a broad field like physics.

Table 1. Citation environment for each journal

¨Seed journal¨	***Number journals (citation environment)***
Scientometrics	271
Physical Review Letters	979
Journals of Vascular and Interventional Radiology	443
Public Health	433

The next step was to create, for each of the four journals analyzed, a network that contained the journal itself, the journals receiving citations from and giving citations to the seed journal, and all citation connections between these journals. This is what we called the Seed Journal Citation Network. For instance, the Scientometrics Citation Network is a network of 272 journals (nodes) connected by the absolute number of citations given (or received) from one journal to another. But as we have argued in the previous section, the absolute number of citations is size affected. To avoid it the links between journals are normalized based on the L index described in the previous section. Table 2 shows the minimum and maximum value of the L index in each of the four networks.

Table 2. L index values for each Seed journal citation network

¨Seed journal¨ citation network	**L index (min value)**	**L index (max value)**
Scientometrics	0.0001	0.1410
Physical Review Letters	0.0001	0.2358
Journals of Vascular and Interventional Radiology	0.0001	0.1886
Public Health	0.0001	0.1578

Once the seed journal citation network was normalized based on the L index, we measured the importance of each of the journals in the network using a centrality algorithm developed by Kleinberg (1999) and explained above. The algorithm gave for each journal of the network two weights: authority weight and hub weight. The journals could then be sorted based on these two weights. The journals with the highest weights were selected for being shown in the network map. The decision as to how many journals are selected is arbitrary. We can show in the map as many journals as we want from the seed journal ego network. When we work with the Netdraw program for the visualization of the maps, we can always zoom in or out to get a better view of the journals involved.

Because this is not possible when you make a fixed "image" of the map, we have just selected a "reasonable" amount of journals having the highest weights values based on the hubs and authorities algorithm. In the selection, a journal that has one of the highest hub centrality weights can have also one of the highest centrality weights among the journals in the seed journal citation network. This is a journal considered as an important source of scientific knowledge and an important source of information to look for the most important authority journals among the journals in the citation environment of the given journal.

The maps then show three types of nodes with different shapes. The squares (blue) are the journals with the highest authority weights in the seed journal citation network, the circles (yellow) are the hubs with the highest hubs weights in the seed journal citation network, and the triangles (red) represent the journals that happen to be at the same time in both of the previous selection. The lines (directed edges) show the citation relation between the journals. The direction of the arrow indicates if a journal is cited by (incoming arrow) or if is citing to (outgoing arrow). The thickness of the connecting line reflects the strength of the L index among a pair of journals.

The position of the journals in the map is based in a spring-embedded algorithm included in the software NetDraw. Its effect is to distribute the vertices in a two-dimensional plane with some separation, while attempting to keep connected journals reasonably close together. As de Nooy, Mrvar, and Batagelj (2005) explained, the edges could be imagined as springs "pulling" vertices (journals) together, though never too close. The algorithm pulls vertices to better positions until they reach a state of equilibrium. In the network journal maps, this layout means that journals that are linked or that have links in common will be closer in the map. It is important to consider though that all the journals are appearing on the map because they have been cited by the seed journal. But in the map, we are considering the strongest citation links (based on the L Index) between the journals selected (25% of the links in the map are taken into account). The program used for visualizing the network maps is NetDraw (Borgatti, 2002).

Scientometrics

Figure 1 shows how Scientometrics is between two groups of journals. One is related to information science and technology journals (right-upper part of the network map) and the other with journals related to research, development, and innovation studies, especially from the management perspective (right part of the network map). It is striking to notice this clear gap between, on the one hand, the scientometrics/library and information science community

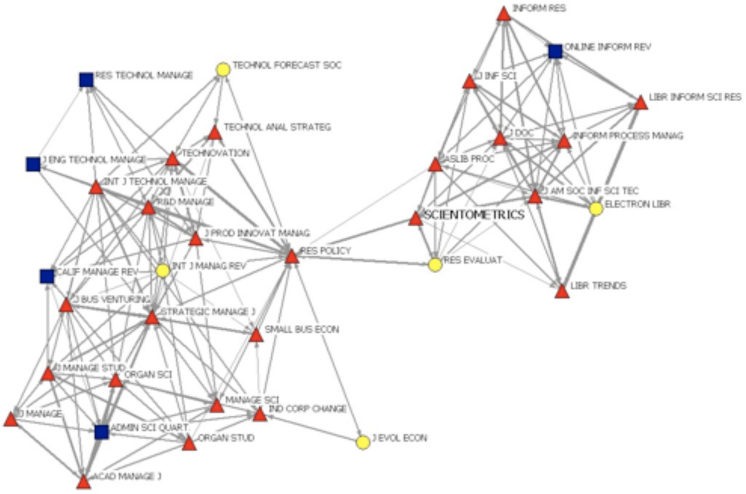

Figure 1. Mapping of the citation environment of Scientometrics (2006) (L index>0.0163)

Squares (blue) Journals with the highest authority centrality weights
Circles(yellow)Journals with the highest hub centrality weights
Triangles(red)-Journals that have at the same time the highest authority and hub centrality weights

Physical Review Letters

Figure 2 shows the central position of Physical Review Letters as well as its status as a hub and authority. Physical Review Letters is first and foremost surrounded by three ‘general’ or multidisciplinary journals. Phtysical Review B is a general physics journals, while Nature and Science are general science journals. Around this first lay, we notice in the network map also the broad coverage of this journal given its strong connection with journals related with physical sub-disciplines such as astrophysics; elementary particles and fields; nuclear physics; atomic, molecular, and optical physics; nonlinear dynamics, fluid dynamics, classical optics; plasma and beam physics; condensed matter; and soft-matter, biological, and interdisciplinary physics.

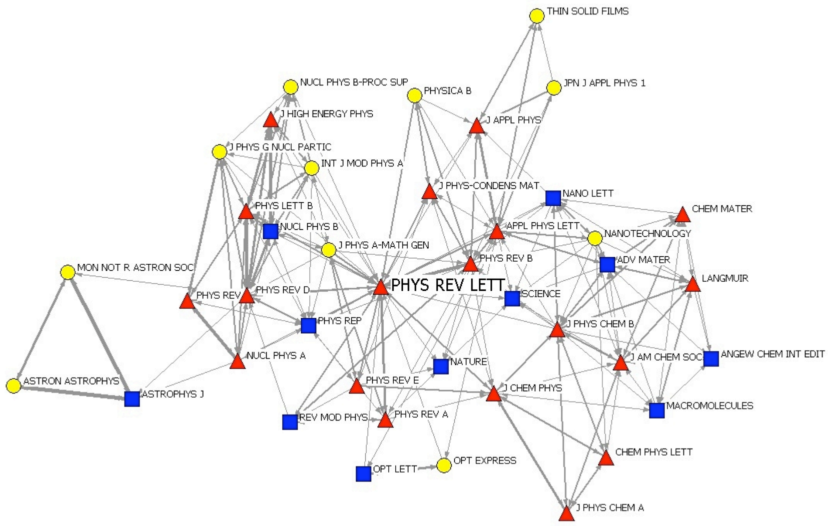

Figure 2. Mapping of the citation environment of Physics Review Letters (2006) (L index>0.0131)

Squares (blue) Journals with the highest authority centrality weights-
Circles(yellow) Journals with the highest hub centrality weights
Triangles(red)-Journals that have at the same time the highest authority and hub centrality weights

Journal of Vascular and Interventional Radiology (JVIR)

Figure 3 shows the position of this journal between journals related with cardiology and vascular surgery in the upper right side of the graph, and journals related with radiology, neurosurgery, and urology on the other side of the graph. JVIR clearly appears as a hub and authority journal.

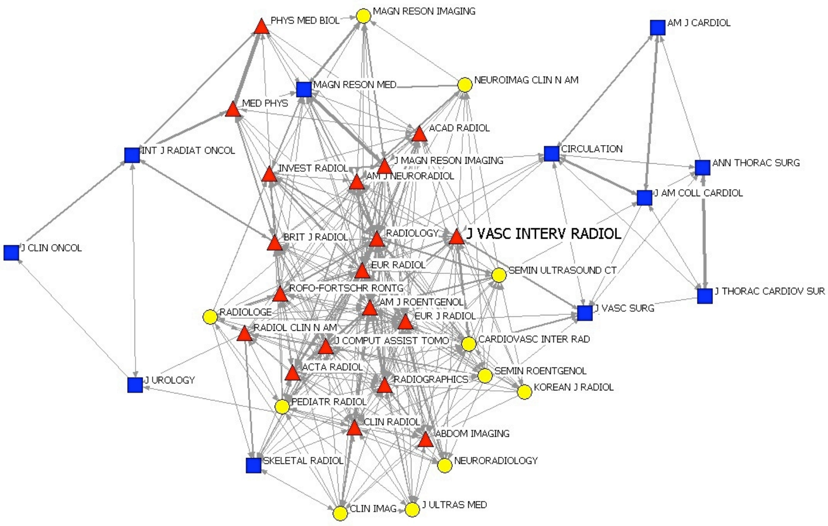

Figure 3. Mapping of the the citation environment of JVIR (2006)
(L index>0.0050)

Squares (blue) Journals with the highest authority centrality weights
Circles(yellow) Journals with the highest hub centrality weights
Triangles(red)-Journals that have at the same time the highest authority and hub centrality weights

Public Health

Public Health is a journal aiming at all public health practitioners and researchers and those who manage public health services and systems. Figure 4 shows its citation network map. Public Health is surrounded by other journals related with public health but none of them have a central position. It is considered a hub spreading the knowledge from journals that are about public sanitary problems as: drugs and addictions, sexual transmittable diseases, obesity, mental health, epidemics, health law and policy.

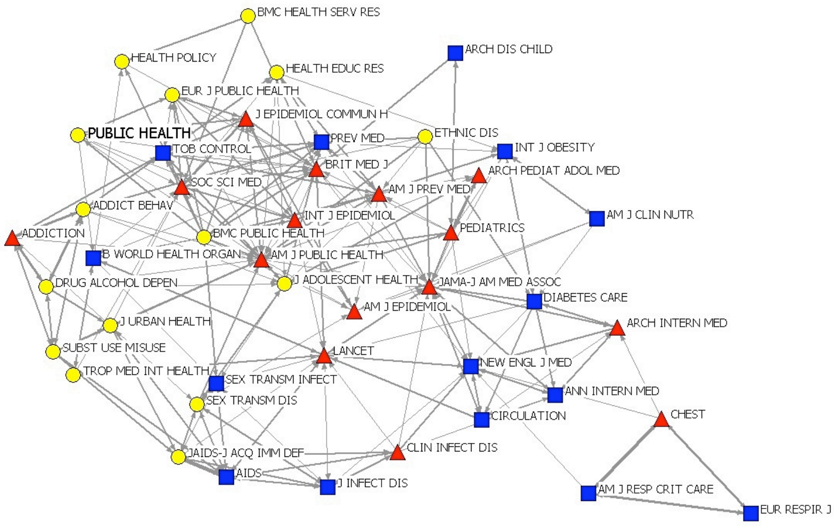

Figure 4. Mapping of the citation environment of Public Health (2006) (L index>0.0089)

Squares (blue) Journals with the highest authority centrality weights
Circles(yellow)Journals with the highest hub centrality weights
Triangles(red)-Journals that have at the same time the highest authority and hub centrality weights

7.5 Conclusions and Follow-Up Research

The method presented here should be considered as the starting point toward a complete methodology for analyzing the citation environment of a journal. From the perspective of bibliometrics and journal performance, our goal was to be able to provide a quick but nevertheless comprehensive overview of the most important journals for a given journal, in terms of citations given and received. The method presented here allows us in a few steps to find the prominent journals, based on a centrality network measure, related to a seed journal. First, we selected the journals in the citation environment of the seed journal. Second, we normalized the citation links inside the seed journal citation network based on an index that considers the importance of the number of citations given from one journal to another in relation to the total number of citations given for the journal and the total number of citations received by the other journal. Then, we apply a centrality algorithm to determine the importance of the journals in the seed journal citation network. From there on we decide how many of these journals we want to represent in a network map. The decision as to how many journals from the seed journal citation network to include in the map is quite arbitrary though.

We are currently working on a further development of the method presented here through dynamic animation of the network map based on time series of the seed journal network data. The objective is a better understanding of the development through time of the seed journal based on its citation relations. Furthermore, we intend to go a step further in measuring the composition and structure of the seed journal citation network. We are interested in studying how bibliometrically related journals form and evolve embedded in the dynamic system of the seed journal citation network. Measures like homophily (Scott, 2000; Wellman, 1993) can help us to determine if journals that have common bibliometric characteristics (such as journal impact measures, degree of international cooperation, degree of journal-to-journal self citations, etc.) stick together. The study of this phenomenon has also been called "assortative mixing in networks" (Newman, 2002), in which the probability of two nodes being connected by an edge depends on specific similarity properties of the nodes. Another measure called homogeneity can determine whether the seed journal's alters are all alike. We can also analyze the structure of the seed journal (the journals to which the seed journal is connected to) citation network with measures like brokerage and density, which measure whether the seed journal connects otherwise unconnected journals.

In summary, the method and results presented here should be considered a starting point for developing a comprehensive methodology to identify from a dynamic perspective the citation environment of a journal.

References

Albert, R., & Barabási, A.L. (2002). Statistical mechanics of complex networks. Review of Modern Physics, 74, 47–97.

Bassecoulard, E., & Zitt, M. (1999). Indicators in a research institute: A multi-level classification of journals. Scientometrics, 44, 323-345.

Batagelj, V., & Mrvar, A. (2006). Pajek: Program Package for Large Network Analysis, University of Ljubljana, Slovenia. Retrieved via: http://vlado.fmf.uni-lj.si/pub/networks/pajek/.

Bergstrom, C.T., West, J.D., & Wiseman, M.A. (2008). The eigenfactor metrics. Journal of Neurosciences, 28(45),11433-11434.

Bollen, J., Rodriguez, M.A., & Van de Sompel, H. (2006). Journal status. Scientometrics, 69 (3), 669-687.

Borgatti, S. P., (2002). Netdraw: Network Visualization Softwres. Harvard: Analytic Technologies. Retrieved via: http://www.analytictech.com/downloadnd.htm

Börner, K., Chen, C., & Boyack, K.W. (2003). Visualizing knowledge domains. Annual Review of Information Science, 37, 179–255.

Börner, K., Maru, J.T., & Goldstone, R.L. (2004). The simultaneous evolution of author and paper networks. Proceedings of the National Academy of Sciences of the USA, 101, 5266–5273.

Bornmann, L., Leydesdorff, L., & Marx, W. (2007). Citation environment of Angewandte Chemie. CHIMIA, 61(3), 104–109.

Brandes, U., & Willhalm, T. (2002). Visualization of bibliographic networks with a reshaped landscape metaphor. Joint Eurographics-IEEE TCVG Symposium on Visualization, D. Ebert, P. Brunet, I. Navazo (Editors). Retrieved via: http://algo.fmi.uni-passau.de/~brandes/publications/bw-vbnr1-02.pdf

Brin, S. & Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems, 30, 107-117.

Calero, C., Buter, R., Cabello, C., and Noyons E. (2006). How to identify research groups using publication analysis: An example in the field of nanotechnology. Scientometrics, 66(2), 365–376.

Chen, C.(1999). Visualising semantic spaces and author co-citation networks in digital libraries. Information Processing & Management, 35, 401-420.

De Nooy, W., Mrvar, A. & Batagelj, V. (2005). Exploratory Social Network Analysis with Pajek. New York: Cambridge University Press.

De Solla Price, D.J. (1965). Networks of scientific papers. Science, 149(3683), 510-515.

Dorogovtsev, S.N., & Mendes, J.F.F. (2002). Evolution of networks. Advances in Physics, 51, 1079–1187.

Garfield, E. (1972). Citation analysis as a tool in journal evaluation. Science, 178, 471-479.

Geller, N.L. (1978). On the Citation Influence Methodology of Pinski and Narin. Information Processing & Management, 14, 93-95.

Hanneman, R.A. & Riddle, M. (2005). Introduction to social network methods. Riverside, CA: University of California, Riverside (published in digital form at http://faculty.ucr.edu/~hanneman/)

He, C., & Pao, M.L. (1986). A discipline-specific journal selection algorithm. Information Processing & Management, 22(5), 405–416.

Kleinberg, J.M. (1999). Authoritatives sources in a hyperlinked environment. Journal of the Association for Computing Machinery, 46 (5), 604-632.

Leydesdorff, L. (1986). The development of frames of references. Scientometrics,

9(3-4), 103–125.

Leydesdorf, L. (1994). The generation of aggregated journal-journal citation maps on the basis of the CD-ROM version of the Science Citation Index. Scientometrics, 31 (1), 59–84.

Leydesdorff, L. (2007). Visualization of the Citation Impact Environments of cientific Journals: An Online Mapping Exercise. Journal of the American Society for Information Science and Technology, 58 (1), 25-38.

Menczer, F. (2004). Correlated topologies in citation networks and the Web. The European Physical Journal, B38, 211–221.

Moed, H.F., de Bruin, R.E., & van Leeuwen, T.N. (1995). New bibliometric tools for the assessment of National Research Performance—Database description, overview of indicators and first applications, Scientometrics, 33 (3), 381–422.

Narin F. & Carpenter, M.P. (1973). Clustering of scientific journals. Journal of the American Society for Information Science and Technology, 24, 425–435.

Narin, F., Carpenter, M.P., & Berlt, N. (1972). Interrelationships of scientific journals. Journal of the American Society for Information Science and Technology, 23, 323–331.

Narin F., Hamilton, K.S., & Olivasto, D. (2000). The development of science indicators in the United States. In: B. Cronin & H. B. Atkins (Eds). The

Web of Knowledge: A Festschrift in Honor of Eugene Garfield (pp. 337–360). Medford, NJ: Information Today.

Nederhof, A.J. (2006). Bibliometric monitoring of research performance in the social sciences and the humanities: A review. Scientometrics, 66, 81-100.

Newman, M.E.J. (2002). Assortative mixing in networks. Physical. Review Letters 89, article no. 208701.

Newman, M.E.J. (2010). Networks: An Introduction. Oxford University Press.

Newman, M.E.J., Barabási, A., and Watts D. (2006). The Structure and Dynamics of Networks. Princeton University Press.

Noyons, E.C.M., Moed, H.F., & Luwel, M. (1999). Combining mapping and citation analysis for evaluative bibliometric purposes: A bibliometric study. Journal of the American Society for Information Science, 50, 115-131.

Page, L., Brin, S., Motwani, R. & Winograd, T. (1998). The PageRank Citation Ranking: Bringing Order to the Web (Tech. Rep.) Standford Digital Library Technologies Project.

Pinski, G. & Narin, F. (1976). Citation Influence For Journal Aggregates of Scientific Publications: Theory, with Application to the Literature of Physics. Information Processing & Management, 12, 297-312.

Pillai, S.U., Suel, T., & Cha, S.H. (2005) The Perron-Frobenius theorem: some of its applications. IEEE Signal Processing Magazine, 22(2), 62–75.

Price, D.J. de S. (1965). Network of scientific papers. Science, 149(3683),510-515.

Pudovkin, A.I. (1993). Citation relationships among marine biology journals and those in related fields. Marine Ecology Progress Series, 100, 207-209.

Pudovkin, A.I., & Fuseler, E.A. (1995). Indices of journal citation relatedness and citation relationships among aquatic biology journals. Scientometrics, 32, 227-236.

Pudovkin, A.I., & Garfield, E. (2002). Algorithmic Procedure for Finding Semantically Related Journals. Journal of the American Society for Information Science and Technology, 53(13), 1113-1119.

Salton, G., & McGill, M.J. (1983). Introduction to modern information retrieval. Auckland, New Zealand: McGraw-Hill.

Scharnhorst, A., & Thewall, M. (2005). Citation and hyperlink networks. Current Science, 89(9), 1518-1524.

Scott, J. (2000). Social Network Analysis: A Handbook. Newbury Park, CA: Sage Publications.

Small, H. (1999). A passage through science: Crossing disciplinary boundaries. Library Trends, 48, 72-108.

Van Raan, A.F.J. (2005). Reference-based publication networks with episodic memories. Scientometrics, 62(1), 549–566.

Van Raan, A.F.J. (2008). Scaling rules in the science system: Influence of field-specific citation characteristics on the impact of research groups. Journal of the American Society for Information Science and Technology, 59 (4), 565-576.

Wellman, B. (1993). An egocentric network tale. Social Networks 15, 423-436.

Zhou, P. & Leydesdorff, L. (2007). The citation impacts and citation environments of Chinese journals in mathematics. Scientometrics 72(2), 185–200.