

# Visualizing Scholarly Impact Over Time

Jason Portenoy and Muhammad Raza Khan | CSE 512 Final Project

## Literature Survey

Visualization of citation networks has been gaining a lot of attention from researchers as it can help in identifying important research papers in different fields, identifying scholarly communities and assessing the impact of disciplines and other fields. [1], [4] and [6] are some of the research papers that deal with this topic using different datasets. In this research project, we intend to analyze a related question that what is the impact of a particular author or research paper on the scholarly literature within and across disciplines, and how can we convey this information in a compelling narrative. This question has not been addressed with the same attention and detail as compared to the other questions related to the Citations Networks Analysis. Though there are ways and metrics to assess the influence of the authors on the field and across the fields, visualization we build needs to be sure to express the importance of each nodes and edge in the network effectively. As such the decisions related to the visualizations become very important. For example, if every citation of a paper is shown on the network then it can result in a cluttered map like the one shown in [3].

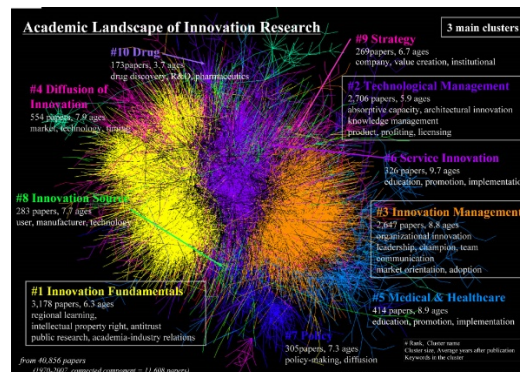


Figure 1 Basic Visualization of Citation Networks [3]

Rosvall and Bergstrom have tried to identify the community structure in the citation networks [1], mapping the citations from one discipline to another and within the discipline. However, they are different from our project in the sense that their focus is on the influence of the citations of the whole field rather than the individual authors. To some extent, the authors' visualizations may be considered to be the macro level visualizations while our aim in this project is to come up with the visualizations that can show the influence of the individual authors within and across fields at the micro level.

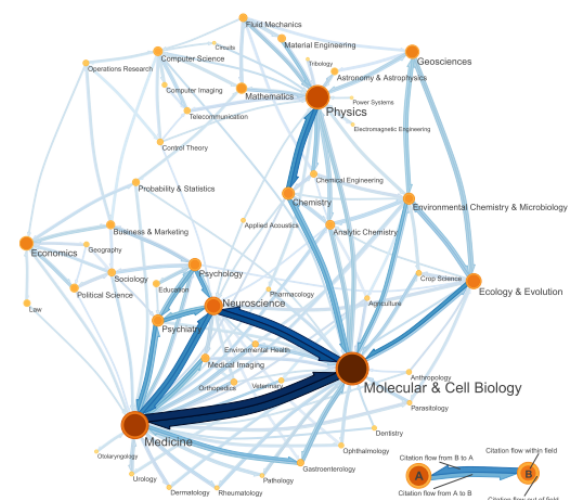


Figure 2 A Map of Science based on Citation Networks [1]

Some of the visualizations decisions used by Rosvall and Bergstrom are visually very intuitive. For example, node size indicates the citation flow with in the network while the ring color width indicates the citation flow outside the networks. We intend to use some of these concepts in our work as well. An extension of the research work in [1], West et al. have tried to evaluate the influence of the authors, institutions and countries on the fields in [2] but their work is more inclined towards the ways and metrics to evaluate the scholarly influence rather than the visualization of the influence. In this project, we want to extend the work in [1] and [2] and visualize the impact of the authors on scholarly literature. Dunne et al have used a force-directed approach to layout papers such that similar papers are close to each other [4] (see figure 3). The papers are shown as rounded rectangles with color coding by their statistical rankings and the spline arrows showing their citations. Communities within the citation networks are shown in the convex hulls. The concept of the Transitive reduction of citations mentioned in [5] can be used to make sure that the edges in the citation network that are not carrying a lot of information can be eliminated.

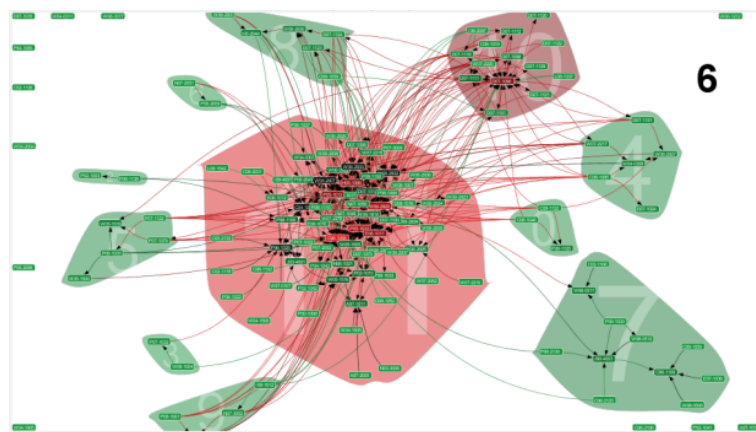


Figure 3 Citation Network Statistics and Visualizations [4]

## References

- [1] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1118–1123, Jan. 2008.
- [2] J. D. West, M. C. Jensen, R. J. Dandrea, G. Gordon, and C. T. Bergstrom, "Author-Level Eigenfactor Metrics: Evaluating the Influence of Authors, Institutions and Countries within the SSRN community," Feb. 2012.
- [3] "kajikawamap.jpg" [Online]. Available: <http://ip-science.thomsonreuters.jp/images/iifl/kajikawamap.jpg>. [Accessed: 20-May-2015].
- [4] C. Dunne, B. Shneiderman, R. Gove, J. Klavans, and B. Dorr, "Rapid Understanding of Scientific Paper Collections: Integrating Statistics, Text Analytics, and Visualization," *J. Am. Soc. Inf. Sci. Technol.*, vol. 63, no. 12, pp. 2351–2369, Dec. 2012.
- [5] N. J. van Eck and L. Waltman, "CitNetExplorer: A new software tool for analyzing and visualizing citation networks," *arXiv:1404.5322 [physics]*, Apr. 2014.
- [6] K. W. Boyack and K. Börner, "Indicator-assisted evaluation and funding of research: Visualizing the influence of grants on the number and citation counts of research papers," *J. Am. Soc. Inf. Sci.*, vol. 54, no. 5, pp. 447–461, Mar. 2003.

## Project Plan

Our initial visualization shows the growth of the network over time (see slides), but does not yet tell a compelling narrative of the growth of influence over time. Next steps involve a lot of exploration of different techniques to make the visualization more intuitive and tell the story effectively. This will include:

- Using clusters to aggregate nodes, which will make the visualization simpler while reducing computational load. We will experiment with different levels of clustering based on distance from the center node.
- Using different encoding channels---color, placement, size of nodes, edge strength---to represent different aspects of the data.
- Experimenting with edge bundling/edge reduction techniques.
- Other narrative techniques to provide more context.

Our narrative will follow the “martini glass” approach of having a guided, animated visualization to start, and then letting the user explore with some degree of interactivity (such as details on demand). We would also like to experiment with interactive ways of showing influence, such as removing the center node and watching how the network changes.