CSE 512 – Data Visualization Spring 2015

Citation Networks http://cse512-15s.github.io/fp-jporteno-mraza/

Jason Portenoy Muhammad Raza Khan {jporteno, mraza}@uw.edu

Problem: The influence of a research paper can be judged by the number of citations it receives, but there are other metrics to consider. Representing citations between papers, a node-link graph can provide a richer picture, but these types of diagrams tend to be cluttered and overwhelming.

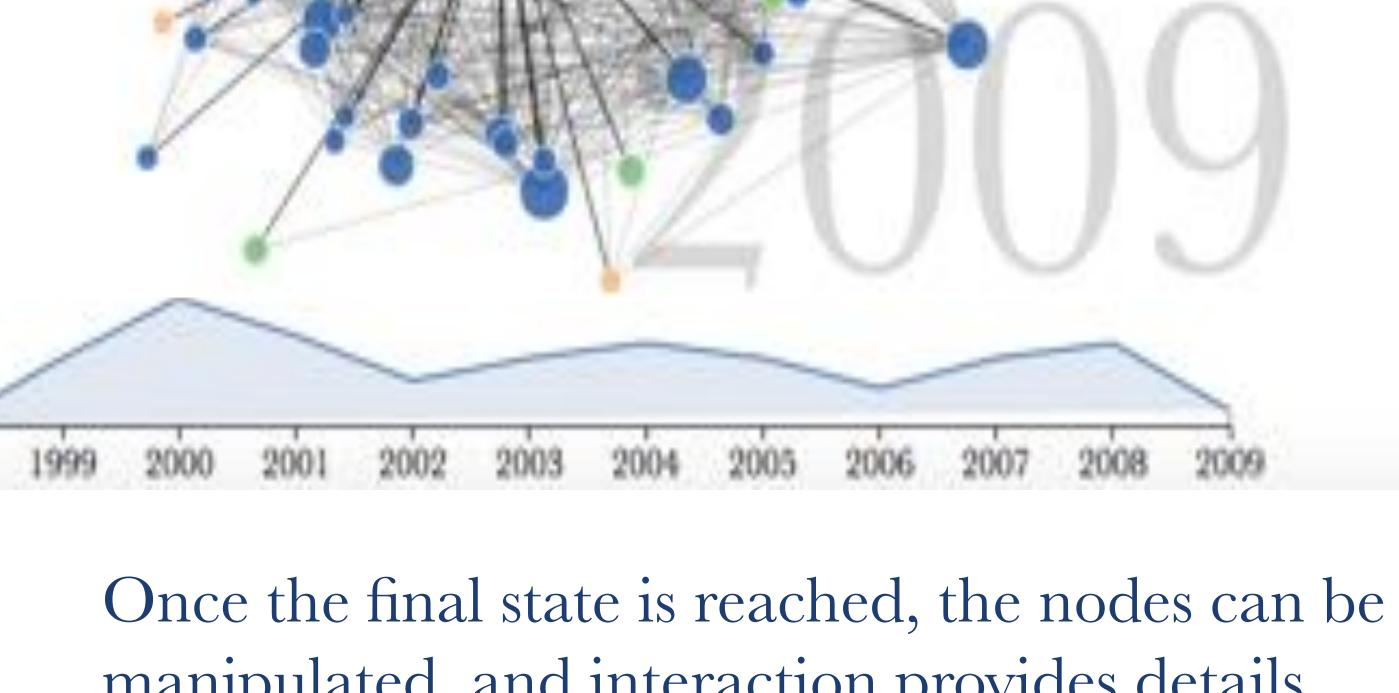
**Approach:** By walking the viewer through the construction of the network year by year, our visualization provides a novel and engaging way of seeing the impact that a research paper has had over time. Initially, the only node shown is the paper of interest. As the year advances, new nodes appear and send out links to the center node and each other, building an ego network. At crucial points, such as a citation from another very influential paper, the animation pauses and an annotation appears to point out the event. Influence of papers are measured using the Eigenfactor [1], and a community mapping algorithm called Infomap [2] detects boundaries between different research fields; these are encoded using node size and color, repsectively. A line chart showing the number of citations per year tracks along underneath the network diagram. When the animation is complete, interaction is opened up and the viewer can get details about the papers in the network.

The initial display shows only the paper of interest, with an annotation giving the paper title and other information.

> Ctf7p is essential for sister chromatid cohesion and links mitotic chromosome structure to the DNA replication machinery was published in the journal Gener & Development in 1999.

Future Work: Experimentation and iterative development will refine the final design, with a more intuitive and meaningful final view that the animation will build to. More annotations, informed by a set of heuristics to identify key points in the development of the network, will provide a richer experience. The framework will be generalized, and real-time database queries will allow for viewing multiple papers in succession. It may also be useful to be able to compare different papers side by side, and to plot multiple papers by the same author. Finally, formal user testing will allow us to validate the design and identify other areas for improvement.

As we move forward in time, new papers are published that cite the center paper. The network diagram builds itself, showing its development year-by-year. The size of nodes shows the relative influence (Eigenfactor value) of that paper, and colors represent different fields of research. A line chart shows both the current year and the number of citations per year.



manipulated, and interaction provides details about individual papers in the network.



<sup>[1]</sup> Bergstrom, C. T., West, J. D., & Wiseman, M. A. (2008). The Eigenfactor<sup>TM</sup> metrics. The Journal of Neuroscience, 28(45), 11433-11434.

<sup>[2]</sup> Rosvall M. & Bergstrom, C. T. (2011). Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems.  $PLoS \ ONE, 6(4)$ .