Literature

Bibliometrics is a method of quantitive analysis of scientific publications. Many current techniques in bibliometrics are focused on citation analysis, looking at how scholars and organizations cite one another in publications. From this data, we can understand the networks and collaborations of scholars and the development and focuses of scientific inquiry over time.

The focus of the majority of bibliometric tools is to underline the relational aspect of publications. In any popular area of science, there are groundbreaking publications that instigated the origin of the field of study. These pioneer publications are almost always cited in any proceeding publication in that area. As equally important, the most prolific writers of the area are commonly cited. Tools such as PaperLens [1] and CiteVis [2] analyze influential authors and references by analyzing the citation mapping. To visualize these relationships, the connections are often demonstrated using network graphs [1-4]. While these visualizations contain a lot of valuable information, the graphical format is not conducive of locating a specific paper within the network to determine that paper's specific contributions and connections. In general, the visualizations can be a mess of wires and nodes that do not facilitate understanding of the relationships. Chinchilla-Rodríguez [5] slightly remedies the mess by grouping publications according to country and organizes a spherical graph according to overall domestic and international citations. However, this graphical simplification also loses the ability to find connections between individual papers.

Tools such as Citeology [6] and CiteRivers [7] organize the papers temporally. This enables the inclusion of individual points for each paper along the time axis, allowing for further analysis of specific papers. Laying out the publications along the x-axis permits addition of a relational notation along the y-axis. For example, Citeology includes wires that span the y-axis to connect the references and citations between papers. For the intended purpose, Citeology may be one of the most concise yet informational visualization format.

To enhance the amount of information available, graphical interactions can be included [7-8]. Specifically, once the relationships between publications are mapped, we may be interested in reading and knowing specific details about a particular influential paper. Giving access to the full citation information of a chosen paper enhances the usability and usefulness of a graphical visualization.

The previously cited works are primarily focused on analyzing the relationships between papers and authors based on citation analysis. While that objective can greatly facilitate a researcher in delving into the literature of their project, the overarching knowledge and view obtained from this foray may be too narrow. To obtain a broader understanding about the topics being studied in a scientific area, we propose organizing publications based on keywords. By searching a specific keyword and tallying the other associated keywords found in publications, we can determine popular subareas of research over time. This offers a overview of the research activity as a whole rather than concentrate on finer detail interactions. By plotting the keyword popularity over time, we can include interactions to also incorporate the paper-to-paper relationships that most bibliometrics are concerned with.

Project Plan

Milestone 1: Raw Data (5/21)

- Mimi - Understand eSearch api and determine necessary commands to download XML file of related papers with specific keywords
- Ben - Write a perl script to automatically call eSearch function for file download

Milestone 2: Integration (5/23)

- Mimi - Determine how to read XML data with D3
- Ben - Determine how to call perl script in front end app

Milestone 3: Data Storage (5/25)

- Mimi - Organize formatted paper information into data structures for easy calling
- Ben - Create lists of paper data structures for D3 visualization (by date and keyword)

Milestone 4: Visualization (5/28)

- Mimi - Draw line plots based on number of papers per keyword over time. Test with a predetermined set of papers.
- Ben - List all papers for a specific publication date and keyword (for additional information when hovering over a specific point in the line plot). Test with a predetermined set of papers.

Milestone 5: Interaction (6/1)

- Mimi - Add text boxes and buttons to visualization for arbitrary keyword search
- Ben - Connect button interaction and backend to update display according to new search terms.

References

1) B. Lee, M. Czerwinski, G. Robertson and B. B. Bederson, Understanding eight years of infovis conferences using PaperLens. In Proc. of the IEEE Symposium on Information Visualization,page 216.3, Washington, DC, USA, 2004. IEEE Computer Society.
2) J. Stasko, J. Choo, Y. Han, M. Hu, H. Pileggi, R. Sadanaand and C. D. Stolper, "CiteVis: Exploring conference paper citation data visually", Posters of IEEE InfoVis, 2013.
3) H. Small, “Visualizing science by citation mapping”, Journal of the American Society for Information Science, 50(9), 799–813.
4) P. Glenisson, W. Glänzel, F. Janssens , B. De Moor, “Combining full text and bibliometric information in mapping scientific disciplines”, Information

Processing and Management: an International Journal, v.41 n.6, p.1548-1572, December 2005.
5) Z. Chinchilla-Rodríguez, M. Benavent-Pérez, F. de Moya-Anegón, S. Miguel, International collaboration in Medical Research in Latin America and the Caribbean (2003–2007), Journal of the American Society for Information Science and Technology, v.63 n.11, p.2223-2238, November 2012.
6) J. Matejka, T. Grossman, and G. Fitzmaurice. Citeology: visualizing paper genealogy. Proceedings of ACM CHI '12 Extended Abstracts, pages 181-190. ACM, May 2012.
7) F. Heimerl, Q. Han, S. Koch and T. Ertl, "CiteRivers: Visual analytics of citation patterns", IEEE Transactions on Visualization and Computer Graphics, vol. 22, no. 01, 2016.
8) P.C. Wong, B. Hetzler, C. Posse, M. Whiting, S. Havre, N. Cramer, A. Shah, M. Singhal, A. Turner and J. Thomas, &ldquo,IN-SPIRE Infovis 2004 Contest Entry,&rdquo, Proc. IEEE Symp. Information Visualization, Oct. 2004.