GeoSocialVis: Visualizing Geosocial Academic Co-Authorship Networks by Balancing Topology- and Geography- Based Layouts

David Saffo 🆔*   Michail Schwab 🆔   Michelle A. Borkin   Cody Dunne 🆔
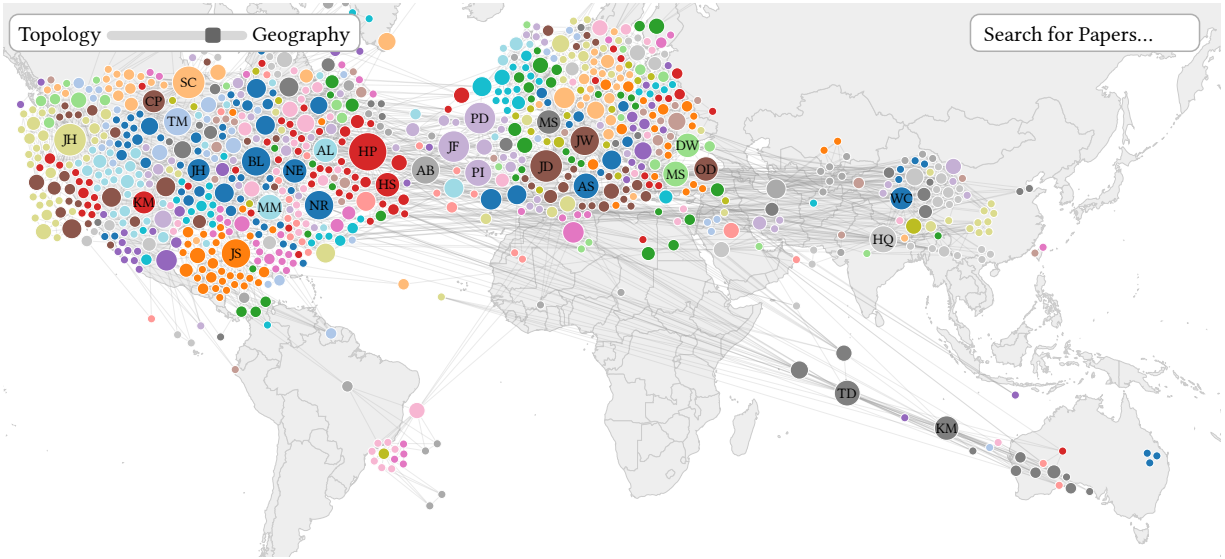
Northeastern University

Figure 1: GeoSocialVis showing the IEEE Information Visualization (InfoVis) 2008–2018 geosocial co-authorship network. The visualization consists of the geosocial network visualization, the topology and geography balance slider, and the paper search panel. Node color encodes the author's most recent affiliation, size encodes paper count, and larger nodes are labeled with author initials.

Index Terms: Human-centered computing—Visualization—Visualization techniques—Graph drawings—Geographic visualization

1 INTRODUCTION

In many domains, it is important to understand both the topology and geography of a network. E.g., for researchers it is important to understand the topics in a field, where this research is taking place, and which researchers collaborate. Geography plays a key role in collaboration as it is more feasible with physical proximity. Researchers are also likely to continue existing collaborations. The interplay between topology and geography is of particular interest. Within-institution collaborations are different in nature from far-distance collaborations, as long-distance collaborations often occur less out of convenience but for more complementary expertise. These long-distance collaborations can connect otherwise separate social groups. However, existing approaches either focus on the geospatial location of researchers, or on the social aspect of collaborations alone.

We present GeoSocialVis, an interactive visualization tool for scientometrics analyses with a focus on displaying the geosocial co-authorship network. GeoSocialVis uses a novel force layout that strikes a user-defined balance between showing network topology and the geographic locations of the authors. Users can explore relevant publications, researchers, groups, and institutions by adjusting the balance between topology and geography, searching by keywords, and with details-on-demand. We demonstrate the utility of GeoSocialVis with a case study with IEEE Information Visualization 2008–2018 publication data. Source code and an interactive demo are available online at dsaffo.github.io/GeoSocialVis

*Corresponding author. E-mails: [saffo.d | schwab.m]@husky.neu.edu, [m.borkin | c.dunne]@northeastern.edu

2 RELATED WORK AND BACKGROUND

Network layout (graph drawing) algorithms for node-link visualization generally focus on leveraging spatial position to encode network topology, e.g. the Fruchterman-Reingold layout [6]. This is contrary to the goal of cartography which uses spatial position to encode geographic location. To show both geography and connectivity, often a node-link visualization is superimposed on a map and/or the map is morphed to display connectivity as in VoroGraph [3] or constrained by boundaries as in IPSep-CoLa [4]. We present a network layout algorithm that combines both of these features into a single 2D spatial encoding. For background on co-authorship network visualization for scientometrics, we refer the reader to the extensive summary by Federico et al. [5]. Here we only mention Chinchilla-Rodríguez et al. [2], who visualize co-authorship networks with egocentric node-link visualizations aggregated by a selected attribute, e.g. by country. This can show high-level geographic relationships but not individual authors and collaborations.

3 VISUALIZATION AND INTERACTION DESIGN

GeoSocialVis is an open source web-based visualization built with JavaScript and D3.js [1]. GeoSocialVis allows users to search for publications while exploring their authors and affiliations.

Co-authorship node-link visualization—The main interface of GeoSocialVis shows each author as a separate node with edges between them showing co-authorship of one or more papers. The network is laid out atop a clipped Mercator projection. Nodes are colored categorically based the authors' most recent affiliation and sized proportionally to their number of publications. GeoSocialVis currently includes the co-authorship network of InfoVis conference papers from 2008–2018 extracted from the VIS publication dataset [7]. We manually performed author affiliation deduplication and

geotagging. The network consists of 907 nodes (authors), 2484 edges (co-authorships), and over 500 geotagged affiliations. From the network visualization, users can identify prolific researchers based on node size and explore their co-authors.

Interactivity—GeoSocialVis' interactions enable exploration of the network. A search panel filters papers by titles, authors, topics, or affiliations. Matching papers are listed with their metadata and DOI hyperlinks. Matching authors and their collaborators are highlighted in the visualization by fading non-matching elements. Users can hover or click individual authors to highlight their collaborators and view their papers in the side panel.

Geosocial Network Layout—GeoSocialVis balances showing co-authorship network topology with showing geographic location of authors using a novel force layout. The network layout is determined by two main forces: (1) authors are pulled towards their co-authors and (2) authors are pulled towards the position of their geotagged affiliation on the map. In addition, node repulsion and overlap removal forces are used. A slider at the top allows users to select an inverse balance between the strength of the topology and geography forces. For a purely topological view, as in a traditional network layout, the slider can be set all the way to "Topology". Alternatively, a traditional map can be seen by sliding to "Geography". Settings in between show a mix of topology and geography which can reveal interesting insights, described in Sect. 3.1.

3.1 Application Examples

In this section, we highlight insights that can be obtained thanks to the balance between the geospatial and social force layout and the connection between the network layout and a powerful search.

Balanced View—In general, the balance between the forces highlights aspects of the academic landscape that are otherwise hidden. For example, Jeffrey Heer ("JH") and Tamara Munzner ("TM") work in the Northwest of Northern America, but Fig. 1 displays Tamara Munzner in a much more central location. Through interaction, users can learn that this is due to more local collaborations by Jeffrey Heer, compared to more collaborations on the U.S. East Coast and Europe by Tamara Munzner. In this way, the balanced network is able to differentiate institutions and researchers based on the amount of local and remote collaborations they use to conduct their research. Researchers such as Jeffrey Heer and John Stasko ("JS") are surrounded by researchers of the same color because they work at the same institution, whereas other institutions are more spread out because of more varied long-distance collaborations.

Understanding Research Areas—Users can use GeoSocialVis to investigate the connectivity of research topics. In Fig. 2 we present two searches for the popular InfoVis research topics "perception" and "color" (49 and 42 publications, respectively). We can see that the co-authorship map for "perception" (top) has a large main connected component and few other components. Conversely, the co-authorship map for "color" (bottom) has a less dense main connected component and there are more disconnected components, though the research comes from many of the same groups. Visualizing such queries can help users begin to dig deeper and answer questions such as: "Who should I work with?", "Who are the biggest players in a topic?", and "How fragmented is research on a topic?"

4 CONCLUSION

We have presented GeoSocialVis, a visualization that focuses on the interaction between social connections and geospatial positions. Our case study uses this paradigm to provide insights about the academic landscape of IEEE Information Visualization. Balancing forces to visualize high-dimensional data on a two-dimensional screen is a promising approach, as it allows users to define the importance of different aspects of their data. We look forward to using this approach in more domains, incorporating more data dimensions to the forces, and evaluating the approach in practice.
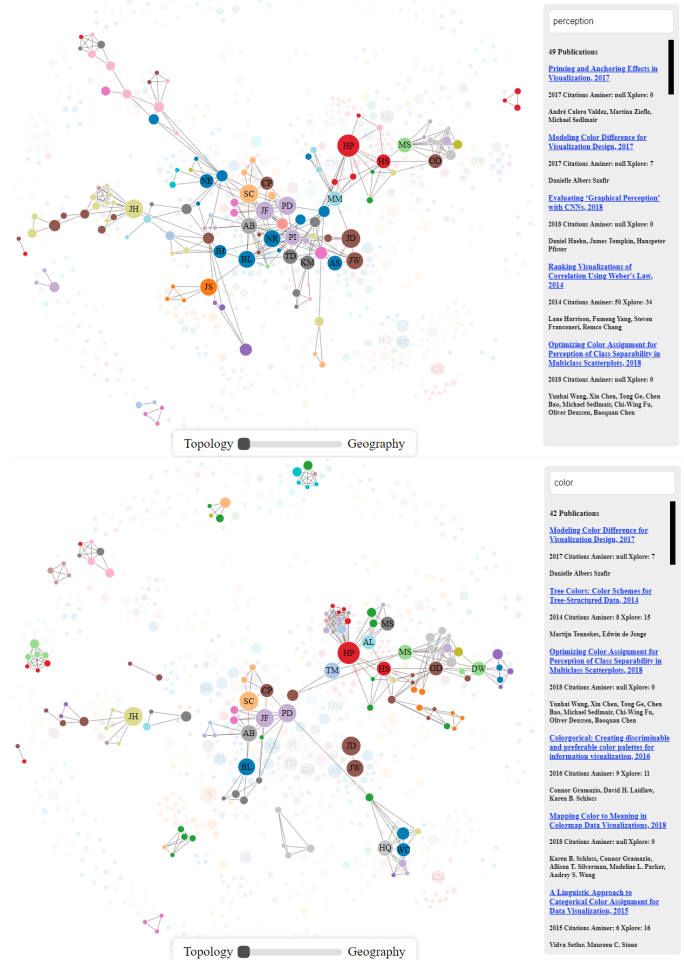

Figure 2: The GeoSocialVis co-authorship network for two searches—"Perception" (top) and "Color" (bottom)—with the force slider set all the way towards "Topology" shows the connections between authors and lists matching papers. Perception research appears to have a more dense research network compared with research on Color.

REFERENCES

[1] M. Bostock, V. Ogievetsky, and J. Heer. D3: Data-driven documents. *TVCG*, 17(12):2301–2309, 2011. doi: 10.1109/TVCG.2011.185

[2] Z. Chinchilla-Rodríguez, B. Vargas-Quesada, Y. Hassan-Montero, A. González-Molina, and F. Moya-Anegóna. New approach to the visualization of international scientific collaboration. *IVS*, 9(4):277–287, 2010. doi: 10.1057/ivs.2009.31

[3] C. Dunne, M. Muller, N. Perra, and M. Martino. VoroGraph: visualization tools for epidemic analysis. In *CHI EA '15*, pp. 255–258, 2015. doi: 10.1145/2702613.2725459

[4] T. Dwyer, Y. Koren, and K. Marriott. IPSep-CoLa: An incremental procedure for separation constraint layout of graphs. *TVCG*, 12:821–8, 2006. doi: 10.1109/TVCG.2006.156

[5] P. Federico, F. Heimerl, S. Koch, and S. Miksch. A survey on visual approaches for analyzing scientific literature and patents. *TVCG*, 23(9):2179–2198, 2017. doi: 10.1109/TVCG.2016.2610422

[6] T. M. J. Fruchterman and E. M. Reingold. Graph drawing by force-directed placement. *Software: Practice and Experience*, 21(11):1129–1164, 1991. doi: 10.1002/spe.4380211102

[7] P. Isenberg, F. Heimerl, S. Koch, T. Isenberg, P. Xu, C. D. Stolper, M. Sedlmair, J. Chen, T. Möller, and J. Stasko. Vispubdata.org: A metadata collection about ieee visualization (vis) publications. *TVCG*, 23(9):2199–2206, 2017. doi: 10.1109/TVCG.2016.2615308