Collaboration Map: Visualizing Temporal Dynamics of Small Group Collaboration

Seongtaek Lim
School of Information
UC Berkeley
Berkeley, CA 94720 USA
stlim@berkeley.edu

Patrick Chiu
FX Palo Alto Laboratory
3174 Porter Drive
Palo Alto, CA 94304 USA
chiu@fxpal.com

Abstract
Collaboration Map (CoMap) is an interactive visualization tool showing temporal changes of small group collaborations. As dynamic entities, collaboration groups have flexible features such as people involved, areas of work, and timings. CoMap shows a graph of collaborations during user-adjustable periods, providing overviews of collaborations’ dynamic features. We demonstrate CoMap with a co-authorship dataset extracted from DBLP to visualize 587 publications by 29 researchers at a laboratory in 25 years.

Author Keywords
Interactive visualization; small group collaboration

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous. See: http://www.acm.org/about/class/1998/

Introduction
It is important that research and development organizations examine their small group collaborations. Analyzing existing collaborations may help them effectively understand the current states of collaboration, in particular, people involved, fields, and timings. Such understanding could suggest future
activities and potential collaborators. However, only few means are currently known and used for this purpose.

Analysis on past small group collaborations within a research and development organization could be a difficult task. One reason is that groups dynamically change over time in nonlinear ways [1]. At a research laboratory, for example, researchers with different skills and interests may collaborate with each other only for a few short-term projects. Thus, a means to analyze the dynamic features of collaboration is necessary.

In this work, we propose Collaboration Map (CoMap), an interactive visualization tool that renders temporal changes of small group collaborations in an organization. This tool may help analyze collaborations in an organization by showing an overview of the collaboration data. We demonstrate CoMap using a dataset of a research organization with 29 members based on their 587 publications over 25 years.

Related Works
Existing graph visualizations on collaboration, such as citation graphs, have several limitations. First, data mining studies focus on theoretical properties of large-scale networks, so that its visualization isn’t applicable for practical use in small- or mid-sized organizations (e.g., [6]). Second, citation graphs focus on specific publication venues (e.g., [3, 4]), but not people in an organization. For the organizational analysis purpose, the visualization should be aligned with people of interest. Third, citation graphs usually give little information on temporal aspects unless it is converted into other forms [4]. Converted to show temporal aspects, it might lose connectivity among vertices, which is generally a useful feature of graphs. Lastly, some information in collaboration graphs might be trivia having little use (e.g., who has the most publications) or serving recreational purposes (e.g., degrees-of-separation) (e.g., [5]).

Collaboration Map
Design
CoMap is an interactive graph where researcher vertices move around clusters of publication venue vertices according to collaboration records during adjustable periods of interest (Figure 1). This relevancy-based arrangement of object clusters is similar to that of Opinion Space suggested by Faridani et al. [2]. The purpose of the tool is to show dynamic features of collaboration: people involved, fields, and timings. We used publication data of a research laboratory extracted from DBLP Computer Science Bibliography [7] including venues with 5 or more papers by the researchers at the laboratory since the year of establishment.

CoMap’s graph has two types of vertices: squares and circles (Figure 2). Squares and circles represent publication venues and researchers respectively. CoMap first sets two-dimensional arrangement of publication venues with force-directed method. All the venue vertices find their positions with right distances from others following the relevancy of each pair, using the number of common properties of them; two venues with many properties in common pull each other, whereas a pair of venues without relevancy pushes away. CoMap uses the number of common terms used in publication titles and venue names at each pair of venues. Like this, CoMap fosters the fixed landscape of collaboration with publication venue clusters.
Once the venue vertices find their positions, CoMap renders researcher vertices. A researcher vertex only appears when the researcher has publication during a specific period. Each venue vertex doesn’t move, but attracts a researcher vertex if the researcher has publication at the venue. If a researcher has multiple publications at two or more venues, the vertex is located in the middle of the venues because their forces of attraction are the same. Like this, CoMap places researcher vertices around relevant venue clusters.

CoMap renders two types of edges. Thick yellow edges represent co-authorship (Figure 1, 2). This type of edges only appears when the researchers coauthored during the selected period. Another type is gray edges that connect researcher vertices with venue vertices at which the researcher has publication. The gray edges help a user understand what venues a researcher vertex connects to. This dimmer visual aid is useful because position of a researcher vertex is ambiguous, especially when either several adjacent venues are near the researcher or the researcher has multiple publications in two or more venues. When a user clicks a researcher vertex, its gray edges turn into purple to highlight the authorship.

CoMap has a timeline with a slider control for period selection. A user may adjust length and position of the time interval in years. When it changes, CoMap updates its graph elements: 1) size and color of each venue vertex, 2) position of each researcher vertex, 3) researcher-to-researcher edges, and 4) researcher-to-venue edges. The size of outer border stands for the number of publications during the whole period, while that of inner box indicates the number within the selected period. A venue with no publications grays out. Researcher vertices move closer to relevant venue vertices where it has publication. Edges connect or disconnect the researchers and venues according to the collaboration record.

**Implementation**

CoMap has three components of a data handler, a web server, and a web application client. The data handler, written in Python, extracts required information from DBLP data in XML format and generates vertex and edge structure in JSON format. The outcome contains publications of the researchers and relevancy scores between venues. The web server running on node.js framework serves the JSON data and application resources. As a simple web server, it serves all the data generated by the data handler to web clients at once. CoMap has a web application client written in JavaScript. When a user visits the web application’s URL, the client retrieves the data and application resources from the server. The interactive force-directed graph visualization uses arbor.js library.
Using CoMap

CoMap provides a user with ways of temporally exploring the collaboration data. When a user adjusts the length and position of the time interval, the graph updates the arrangement of vertices and edges to show an overview of small group collaborations in the organization. The followings are visual features in CoMap and their implications.

- Proximity of researchers indicates their common areas of interest during the period; this suggests potential collaborators within an area.
- Scope of movement of a researcher around venue clusters means frequency of interest shifts; this suggests focused or well-rounded researchers.
- Length of purple edges of a researcher means his level of interdisciplinarity; a researcher with long purple edges might be a potential collaborator who tends to reach out to other fields.
- Length of yellow edges means levels of interdisciplinarity of the group; a group of researchers connected with a larger yellow polygon means a more interdisciplinary group.

Conclusion

We proposed an interactive visualization tool that shows dynamic features of collaboration in an organization. Using this tool gives a sense of how the collaborative groups changed over time.

This work still has room for improvement. To resolve occlusion problem of the vertices can improve the visualization. Also, better layout algorithms for venues and more semantically meaningful properties between venues can improve the cluster representation. In addition, a user study is required to verify the tool’s effectiveness.

Acknowledgements

We thank Don Kimber, Frank Shipman, and Khai Truong for valuable comments on the project.

References