Visualizing *InfoVis* Reaserchers with *ContactTrees*

Arnaud Sallaberry*
University of California at Davis

Kwan-Liu Ma†
University of California at Davis

ABSTRACT

This poster presents an application of *ContactTree*, a new ego-centered visualization design, to charactering the collaborative activities of selected *InfoVis* researchers in terms of their publications as listed in DBLP. The ContactTree visualization, based on a botanical tree metaphor, is designed for studying individuals and comparing their social behavior and relationships with others based on data with rich attributes. As shown, the resulting trees give each of the selected researchers a distinct look, and many interesting patterns reveal themselves. In social science research, the study of people's social contacts and activities is of high interest. ContactTree visualization well complements conventional network visualization, which is better for showing relations among people and activities in a global context.

Index Terms: Computer Graphics [I.3.3]: Picture/Image Generation—

1 INTRODUCTION

The analysis of social networks, which are composed of people (actors) and their relationships (ties), has been the subject of a very dynamic field of study for decades [7]. Nowadays, the development of new information technologies helps to collect large data sets to discover many properties related to the formation of social web or the development of interpersonal relationships. Besides the scientific interest of such studies, the recent success of online social networks, personal such as *Facebook* as well as professionals such as *LinkedIn*, shows how the masses are concerned with creating and maintaining their own network influence. First reserved to sociologists, the study of social networks has quickly attracted other disciplines. The emergence of methods and tools devoted to social network analysis and visualization (*e.g* [1, 2, 3, 5]) shows the popularity of *InfoVis* approaches. Most of these methods consider social networks as a whole relying on a global or complete network approach. As comprehensive as it can be, the global approach fails to capture some aspects about social relationships and the study and comparison of the ego-centered social circle representations allows one to discover trends that global representations fail to highlight [4].

ContactTree is a new ego-centered visualization design that helps to access data, compare persons' interpersonal relationships and make hypothesis about patterns or trends. The design is based on a botanical tree metaphor. The main idea is to use the features of a tree (the structure of its branches, leaves, fruits, colors, *etc.*) to map properties of one's relationships. This design fits well to visualize many aspects of one's social circle. We present here an example involving the collaboration network from 2000 to 2011 of some well-known researchers in information visualization. A technical report including a more complete description of the design and several case studies is also available [6].

*e-mail: asallaberry@ucdavis.edu
†e-mail: ma@cs.ucdavis.edu

2 DESIGN DESCRIPTION

A tree can be seen as a structure composed by a trunk, which is divided into several parts to form main branches, themselves divided into smaller ones, *etc.* Starting from this observation, an intuitive approach consists in representing ties as the smallest branches. According to this, a small branch holds exclusive properties of a tie, while bigger branches represent common properties of the ties represented by the small branches growing up from them. Finally, the aspect of the trunk represents the whole relationships.

As the set of examples proposed in this poster is based on a co-authorship network, each tree focuses on an *InfoVis* researcher and the ties are his co-authors. The starting point of main branches growing up from the trunk can be characterized by two parameters: its position along the trunk and the side of the trunk it is on. Then, we can map on it two properties. The first one is on an ordinal scale (position along the trunk) and the second one is a boolean (right or left side of the trunk). In our examples, co-authors lying on the right side of the tree are persons presently (*i.e.* 2011) working in the same country as researchers while those lying on the left side work in another country (see Figure 1). Y-position of a main branch along the trunk represents the year of the first collaboration during the period, from 2000 (lowest branch) to 2011 (highest branch). These two properties characterize the set of ties represented by the main branches.

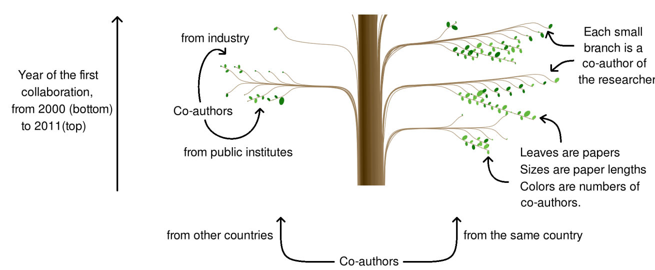


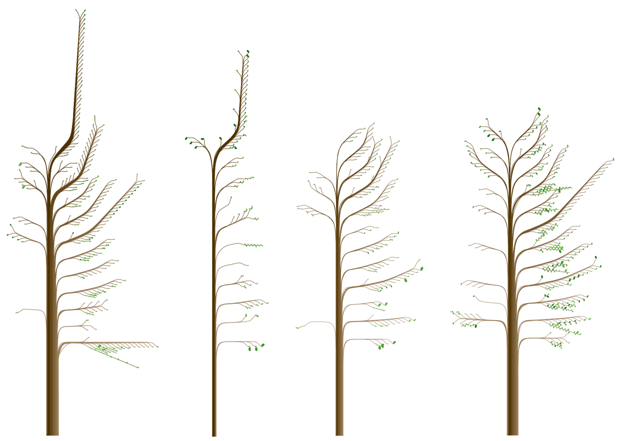
Figure 1: Features of the tree we map properties on.

In the same way, the position along the main branch and the side (above/bellow) where a smallest branch is starting from can be used to represent properties (one ordinal and one boolean). In our examples, positions are not used while small branches growing above (resp. below) main branches represent co-authors from industry (resp. public institutes).

Proceeding the botanical metaphor, leaves need to be added to these branches. They can represent two kinds of properties of ties. Each leaf also holds four features we can map a property on: color, size, position along the branch, side of the branch they are growing from. In our example, each leaf is an article published between 2000 and 2011 and written by the researcher and his co-author represented by the small branch it is growing from (see Figure 1). The size of a leaf is proportional to the number of pages of the article and its darkness represents the number of authors who have collaborated on it.

3 RESULTS

Figure 2 shows the *ContactTrees* of four U.S.A researchers in Information Visualization: Ben Shneiderman, Jeffrey Heer, John Stasko and Kwan-Liu Ma. These *ContactTrees* are very unbalanced: most co-authors lie on the right side of the trunk and on the lower side of the main branches. It means that most collaborators are academic researchers working in the same country as the researcher represented by the tree. This trend is not always true. For example, we see that Ben Schneiderman had many collaborators from industry during the year 2009. This ”anomaly” is mainly due to 2 articles involving many industrial researchers. Another important collaborator who is pointed out by our visualization is Catherine Plaisant. She is represented by the long small branch growing from the lowest main branch. She has written more than 30 papers with Ben Shneiderman during the given period. One last interesting remark on the *ContactTrees* of Figure 2 is related to the distribution of new ties over the years. We see, for example, that in 2004 and 2005, Jeffrey Heer started to work with two new collaborators (Danah Boyd and Maneesh Agrawala, with whom he has published 12 papers). In contrast, the number of his co-authors sharply increased in 2011 with 32 new ties (with up to three papers for one of them). It would be interesting to see the evolution of these collaborations in the future. Maybe 2011 is an important step for Jeffrey Heer's career.



(a) (b) (c) (d)

Figure 2: Four U.S.A researchers: (a) Ben Shneiderman, (b) Jeffrey Heer, (c) John Stasko and (d) Kwan-Liu Ma.

Let's also have a look at two other well-known researchers from Europe: Jark van Wijk, from Netherlands, and Jean-Daniel Fekete, from France (Figure 3(a,b)). As for the previous examples, we see that most of their co-authors come from public institutions. Their *ContactTrees* are much more balanced regarding the country of origin of their ties. This is likely due to the smaller number of team devoted to Information Visualization in their respective countries. However, this observation must be tempered. We see that most of foreign co-authors, as for U.S.A researchers, have been involved in only one or two papers. So, these kind of collaborations are weaker.

Figure 3(c,d) shows two more recent researchers in Information Visualization: Nathalie Henry Riche and Tim Dwyer. They both work in the industry. Looking at the side of the main branches where the small branches start, we see that Nathalie Henry Riche (resp. Tim Dwyer) mostly work with industrial partners since 2008 (resp. 2010). This is because they have joined *Microsoft Research* in 2008 while they were previously in a public institution. Another interesting aspect of their career is shown by the side of the trunk their ties lie on. Nathalie Henry Riche started her career in France, Tim Dwyer in Australia. They both now work in U.S.A. Looking at Tim Dwyer's *ContactTree*, we clearly see that after 2010, most of his new collaborators are in U.S.A. This phenomenon is less salient for Nathalie Henry Riche. We see that she is still having a lot of new ties from another country.

(a) (b) (c) (d)

Figure 3: Two European researchers: (a) Jark van Wijk and (b) Jean-Daniel Fekete. Two more recent researchers: (c) Nathalie Henry Riche and (d) Tim Dwyer, both work in a U.S.A company and both come from another country

4 Conclusion

Our ContactTrees allow people to map ties and contacts based on various priorities and preferences. Facing overly sophisticated network data, whether they are taken from contact diaries, citation records, survey archives, or online social media, researchers and users would benefit greatly by first looking at the key features of a tree. These features capture not only ”ties”, the underlying unit of analysis in most social network analyses, but also ”contacts” among ties. The addition of this key property distinguishes our design from the majority of visualization tools for social networks.

This work is performed by *InfoVis* researchers in collaboration with sociologists. We're currently working on a second dataset, provided by the sociologists, and including individual ties and their contacts for 2 waves of 3 months. While many properties are related to the individuals, their ties and their contacts, more features of the tree are used, including fruits and even birds. We're also exploring different mappings of properties with the same dataset. The technical report [6] presents the current results.

References

[1] M. Baur, M. Benkert, U. Brandes, S. Cornelsen, M. Gaertler, B. Köpf, J. Lerner, and D. Wagner. Visone. In *Graph Drawing*, pages 463–464, 2001.

[2] J. Heer and D. Boyd. Vizster: visualizing online social networks. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis'05)*, pages 32–39, 2005.

[3] N. Henry. *Exploring Large Social Networks with Matrix-Based Representations*. PhD thesis, Université Paris-Sud (France) and University of Sydney (Australia), 2008.

[4] C. Kadushin. *Understanding Social Networks: Theories, Concepts, and Findings*. Oxford University Press, 2012.

[5] A. Perer and B. Shneiderman. Integrating statistics and visualization for exploratory power: From long-term case studies to design guidelines. *IEEE Computer Graphics and Applications*, 29:39–51, 2009.

[6] A. Sallaberry, Y.-C. Fu, H.-C. Ho, and K.-L. Ma. ContactTrees: Ego-centered visualization of social relations. Technical Report CSE-2012-24, June 2012.

[7] J. P. Scott. *Social Network Analysis: A Handbook*. SAGE Publications, 2000.