NetworkCube: Bringing Dynamic Network Visualizations to Domain Scientists

Benjamin Bach, Nathalie Henry Riche, Roland Fernandez, Emmanoulis Giannisakis, Bongshin Lee, Jean-Daniel Fekete

NetworkCube: Bringing Dynamic Network Visualizations to Domain Scientists

Benjamin Bach*
Microsoft Research - Inria Joint Centre

Nathalie Henry Riche†
Microsoft Research

Roland Fernandez‡
Microsoft Research

Emmanoulis Giannisakis§
Inria France

Bongshin Lee¶
Microsoft Research

Jean-Daniel Fekete‖
Inria France

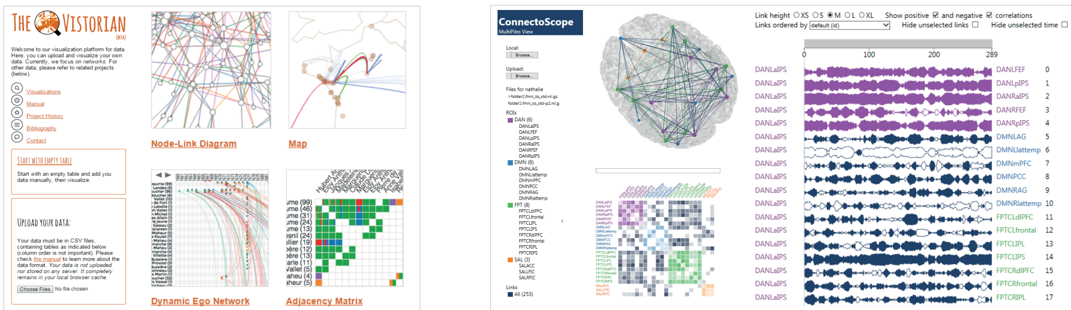

Figure 1: The Vistorian (left), an online web platform to enable social scientists and historians to explore dynamic social networks, and ConnectoScope (right), an online web platform to enable neuroscientist to visualize dynamic functional brain connectivity networks.

Abstract

Network visualizations support research in a range of scientific domains from biology to humanities. This paper presents our effort to create a platform to bridge the gap between domain scientists and visualisation researchers; *NetworkCube* aims in being a fast way to deploy experimental visualizations from research to domain experts analyzing dynamic networks. In turn, InfoVis researchers benefit from studying how their visualizations are used in the wild.

1 Introduction

Research on interactive network visualizations has made tremendous advancements, tackling multivariate and dynamic networks [3], for example. However, as few of them are readily available to end-users and little is known about how well these visualizations support visual exploration in application domains, such as the humanities or neuroscience.

On the one hand, there is this plethora of novel interactive visualization techniques focusing on specific aspects of the data, often tying them to a specific application domain such as biological or social networks. While these interactive visualizations often remain a proof-of-concept or prototype, they can still enable discoveries in the application domain they target [4]. However their applicability may prove hardly generalizable to other domains, tasks, or datasets and they are rarely made available due to their lack of robustness.

On the other hand, a small set of robust network visualizations tools, such as Gephi (`http://gephi.github.io`), NodeXL (`http://nodexl.codeplex.com`), or Cytoscape (`http://www.cytoscape.org`) are available to a larger audience. They provide full-featured well-known techniques but often do not incorporate novel visualizations and visualizations for dynamic networks. A promising approach in this direction is NodeXL, but in general adding novel techniques is hard, as they are numerous, may be specific to an audience or dataset, and require significant engineering time to integrate. The result is often a dead-lock situation in which novel techniques are rarely made available to the larger research community and cannot be proven generalizable because they are not implemented and wide-spread, which prevents end-users from using them to advance research their own field. Thus, these generic tools are rarely leveraged by the InfoVis research community.

We introduce the *NetworkCube*, a light-weight JavaScript framework for the visualization of dynamic networks aiming at providing novel and work-in-progress, yet robust, visualizations to domain experts, while driving advances in visualization research (`https://github.com/benjbach/networkcube`). NetworkCube is motivated by lessons learned from building two interactive network visualizations platforms: The Vistorian and ConnectoScope (Figure 1), issued from collaborations with historians and neuroscientists. The goal is to find the right level of generalizable visualizations that target actual domain problems, e.g., temporal changes, weighted edges, long series of temporal records. While domain experts can benefit from new visualizations in that they may allow them to make new discoveries, infovis researchers can connect to experts using the platform, experiment novel visualizations, and study how users visually explore their data.

2 Visual Exploration of Dynamic Networks

Our collaborations with historians and neuroscientists yielded novel network visualization prototypes (`http://www.visualizingbrainconnectivity.org`) and two platforms; 1) The Vistorian for social networks extracted from history documents, 2) ConnectoScope for functional brain connectivity networks.

The Vistorian: Network analysis and visualization is a relatively recent phenomenon among historians. Visualizations other than node-link diagrams are not yet part of their methodological

*e-mail: benj.bach@gmail.com
†e-mail:nath@microsoft.com
‡e-mail:rfernand@microsoft.com
§email:em.giannisakis@gmail.com
¶email:bongshin@microsoft.com
‖email:jean-daniel.fekete@inria.fr

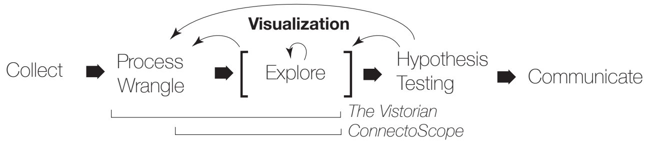

Figure 2: Exploration workflow

corpus, and these are most often used to illustrate findings rather than to explore data. In the Vistorian, users map their manually assembled tables to a network structure, i.e. defining columns for source and target node, edge type and weight; then visualize their data among four visualizations: node-link diagram, adjacency matrix, geographical map with node-link networks, and a novel visualization for dynamic ego-networks.

ConnectoScope: Network analysis in neuroscience is far more advanced [6], yet mostly based on statistics and network measures. This is certainly due to the complexity of data: three-dimensional, temporal, multivariate, multiple data sets, noisy. In a nutshell: this data is hard to visualize. The preprocessing pipeline is rather complex and relies on decisions made by the analyst. The current workflow (Figure 2) does not really integrate exploration, but rather involves forming series of hypotheses and developing and tuning analytical models to validate or reject them. While visualizations often appear in communicating findings, we observed that they can play a crucial role to generate hypotheses prior to labor intensive analysis.

No visualization for dynamic networks is available and tools rarely exhibit interaction capabilities. In ConnectoScope, users upload data in a neuroscience specific format (NIfTI) plus a set of brain region locations. Unlike The Vistorian, ConnectoScope does not support the manual data wrangling pipeline, it automatically extracts the connectivity network with the passed parameters and shows three visualizations: 3D glass brain (and node-link diagram), adjacency matrix, and LinkWave [5], a visualization depicting changes of weighted connections over time. Views are related by brushing and linking, and users can create bookmarks on brain regions (nodes) or connections to be colored across views. We observed that neuroscientists used this coordination mechanisms to learn how to interpret LinkWave.

2.1 Lessons Learned

We noted many commonalities between both projects with respect to prerequisites and expectations, motivating NetworkCube.

[L1] Data wrangling is different per application domain but results in similar graphs, which can be viewed with the same set of common visualizations.

[L2] Researchers in other fields tolerate bugs. Many of their tools are maintained by their community and a few sessions on a prototype can lead to discoveries and save hours of analytical work.

[L3] Most domain experts are novice visualization users. Standard visualizations (e.g. node-link diagrams) can be instrumental to teach more advanced ones via multiple coordinated views.

3 NetworkCube

NetworkCube is our attempt to unify the common parts —data storage and management, visualizations, view updates and events, image export, annotations, search, etc.—for both projects, but keep domain specific routines in separate projects. The main philosophy behind NetworkCube is to provide an extensible set of simple visualization components targeted to specific tasks, data aspects and possibly domains, that can be coordinated together to enable visual exploration. This provides a single code base to maintain and extend, and most important allows to reuse visualizations (L1). Since data formats and formatting needs are specific to each domain, networkcube requires networks to be formatted into a simple pair of node and edge tables. Data is stored exclusively on the client and stored locally until users erase their cache. This prevents users from being worried about uploading sensitive data to a server, or having to run their own server.

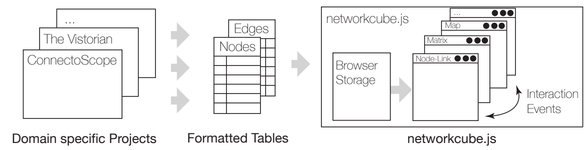

Figure 3: NetworkCube pipeline

NetworkCube is packaged into a single file *networkcube.js* that can be loaded into any web application. NetworkCube comes with the visualizations currently available in The Vistorian and ConnectoScope are reimplemented with NetworkCube, which will be extended with visualizations from current research (L2) (e.g.,[1, 2]). NetworkCube requires only a simple function call to create a visualization for the loaded network in a new browser window. Multiple views can be opened simultaneously and a client-side messenger passing events between them enables full multiple coordinated view experience and easy maintenance of event history. This allows to restore the visualization after a crash as well as log users' interactions. NetworkCube provides a set of standard and familiar visualizations (e.g. node-link diagrams, adjacency matrices) that can help teach users about novel ones as they brush and link elements across views (L3). Connectiong a visualization boils down to obtaining the data as in-memory graph object (one single call), and implementing the desired message event handlers.

4 Research with NetworkCube

As novel components are incorporated into NetworkCube, we believe it may attract more users and lead to advances in different domains. We plan to leverage it for visualization research:

1. Understanding visual exploration in the wild: while our research community has deployed many tools to support data exploration using interactive visualizations, there are only a few studies of long-term data exploration with real users. We believe that instrumenting NetworkCube and making it available to researchers in other domains may yield many such studies.

2. Investigating strategies to teach novel visualizations more effectively: while the envisionned users of networkcube are experts of their data, most remain novices in visualization. We envision that NetworkCube will enable testing different strategies to help people learn novel visualizations, in realistic contexts.

3. Interactive visualizations for communication and scientific storytelling: while much research has been devoted to exploration and analysis, little is known about how interactive visualizations can effectively help communicate findings (storytelling). NetworkCube could potentially shed the light on the entire process from data exploration to visually shared stories.

References

[1] B. Bach, N. H. Riche, T. Dwyer, T. Madhyastha, J.-D. Fekete, and T. Grabowski. Small MultiPiles: Piling Time to Explore Temporal Patterns in Dynamic Networks. *Computer Graphics Forum*, 2015.

[2] B. Bach, C. Shi, N. Heulot, T. Madhayastha, T. Grabowski, and P. Dragicevic. Timecurves: Folding time to visualize patterns of temporal evolution in data. In *IEEE TVCG*, 2015.

[3] A. Kerren, H. C. Purchase, and M. O. Ward. *Towards Multivariate Network Visualization*. Springer, 2014.

[4] N. H. Riche, B. Bach, R. Fernandez, B. Lee, T. Grabowski, and T. Madhayastha. Connectoscope: Interactive data visualization for functional brain connectivity. In *Proc. of Organization of Human Brain Mapping*, 2015. poster.

[5] N. H. Riche, Y. Riche, N. Roussel, S. Carpendale, T. Madhyastha, and T. J. Grabowski. Linkwave: A visual adjacency list for dynamic weighted networks. In *Proc. of Interaction Homme-Machine (IHM)*, pages 113–122, New York, NY, USA, 2014. ACM.

[6] O. Sporns. *Networks of the Brain*. MIT Press, 2010.