SIGCHI Conference Proceedings Format

Yang Chaoyu, Aniket Handa
University of Washington, DUB Group

e-mail address (Blank if Blind Review)
Optional phone number (Blank if Blind Review)

Greg L Nelson, A Conrad Nied
University of Washington, CSE, DUB Group
Address (Blank if Blind Review)
e-mail address (Blank if Blind Review)
Optional phone number (Blank if Blind Review)

ABSTRACT

Understanding dense, time-varying networks lies at the heart of long-standing challenges in scientific domains, especially biology. We construct a focused visualization tool for neurological functional connectivity data. We apply multiple filtering techniques – animation and selecting by inclusion/exclusion - to make displays less overwhelming for general and expert audiences. We also allow filtering with a bootstrapped confidence level and support 2-way interactive visual hypothesis testing, to bring the noise and complexity of the data into context. Our animation display also supports casual use and inquiry into the underlying data for a general audience, which may assist in crowdsourcing visual hypothesis testing.

Author Keywords
Neurology, Neuroscience, Visual Hypothesis Testing, Casual Data Analysis

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
We sought to create a way to visualize this dynamic network flow as an interactive graph and accompanying diagrams. Existing network flows can be beautiful and useful, from Network flow is not new, traffic flow is commonly visualized, as well as many other novelties such as wind, air traffic. Within the neuroscience literature, many groups have designed neural visualizations, and the area has received more attention from visualization researchers more recently. One of the largest limiting factors of all of this data though is that it relies on arbitrary computer programs or hand-tailored videos that support - and never challenge - a designers hypothesis.

CHI 2009, April 4–9, 2009, Boston, Massachusetts, USA.

Many of these visualizations fail to express the data and effectively communicate it. Dense networks have too many edge crossings and it is very difficult for users to disambiguate connections. Time-varying data overwhelm static representations, and end-users have trouble without multiple levels of filtering tools.

Our mission for this project is twofold. We aim to improve on the techniques done by scientists to visualize network flow, and we also want to give the public a tool to explore network flow and evaluate it.

Domain: Neuroscience and Functional
We sought to capture the dynamics of network flow by examining a functional connectivity model of neurological activity. Functional connectivity measures the degree in which regions of the brain communicate or interact as they process certain events. Generally, our approach can be used to flow any time-varying network interactions, but we focused on this dataset to motivate our decisions.

The data comes from brain imaging experiments conducted at Massachusetts General Hospital. Subjects listen to an auditory stimulus to perceive whether they heard an English word or not. Paul A. Luce created lists of these stimuli and grouped them by how many words were just one phoneme (sound-specific letter) different and how often its constituent phonemes were present in the language [1]. While subjects did this task, their neural activity was recorded by electrical and magnetic sensors (EEG and MEG). Neural activation was mapped on to the grey matter surface and clustered into 40 nodes. The waves of cortical activity were used to measure functional connectivity by using a Granger causality model using Kalman filters. The data given to us is the measure of connectivity from the first 600 milliseconds after the subject starts to hear a word, conditioned by different word groups.

The data demonstrating the interaction between nodes is very noisy, due both to the nature of the sensors as well as the multi-layered processing required to convert the readings into associated estimates for regional brain activity. Determining which apparent interactions represent meaningful relationships remains a significant challenge for the field. We explored transforming the data in a few ways to measure a sense of significant connections. As illustrated in Figure X, one transformation of the data results in a time-varying value called the Granger causality index (GCI).

Values greater than zero show signs that an area is dependent on another, values less than show otherwise but their magnitude is, in isolation, meaningless. By taking the cortical activity data, randomly shuffling the values, and putting it back into the causality model, we produced a distribution of GCIs of clearly non-causal data. The distribution of experimental GCIs has a mean of -0.018 and a standard deviation of 0.089 while the null dataset had a mean of 0.003 and a standard deviation of 0.050, with 3.8 million data points in both sets.

Thereby, the majority of the dataset's values, including many of them above zero, can very well be produced by random noise. As a backwards t-test, we filtered the data by rejecting all signs of causality with a GCI below the 95th (0.0733) or 99th (0.1228) percentile of the null-models GCIs. This creates a binary condition (significant evidence of causality or no) for each region to region connection over each time point. We created edges in the dataset to represent

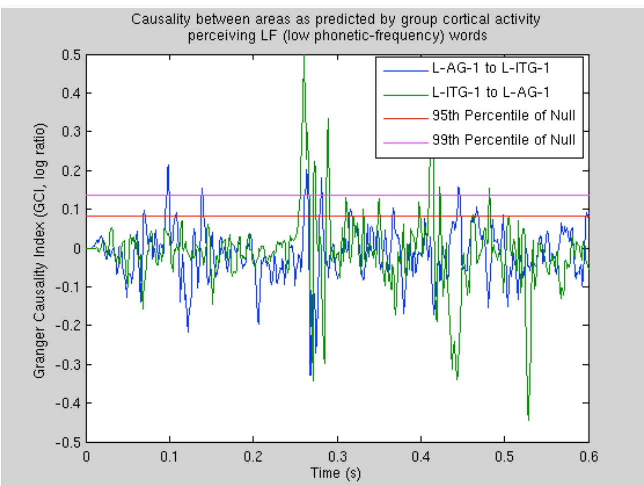

Figure 1. Figure captions should be centered and placed below the figure.

RELATED WORK

Layout

Most prior work focuses on showing data mapped onto the physical brain layout [2,3,4]. This takes advantage of expert knowledge in the domain to provide additional context without labeling or overlays. In neuroscience in particular many groups have their own region enumeration and descriptions, owing to the complexity and flexibility available to interpret gathered data. These techniques tend towards the goals of scientific visualization and focus less on interaction techniques. We considered the opposite side of the design spectrum, starting with a simple circle layout and allowing the user to experiment with different arrangements of nodes.

Complete Datasets

Work on dense, time-varying graphs has applied dotplots and other high-density techniques to attempt to show the full picture of the data. We start from an overview of the data and influence analysis to start small, looking at behavior among small sets of nodes.

Uncertainty and Hypothesis Testing

Prior work focuses on small multiples, instead we show data on the same plot, ask user to mouse over to reveal true data or not.

Prior work does not consider hypothesis testing for dense, time-varying networks. We are also unaware of work in visual hypothesis testing for networks in general. We chose to focus on interactive visual analysis testing and consider the tension between showing and filtering data according to significance levels derived from potentially flawed models, where we might show it to the user unaware of whether it is real or fake data from the beginning of analysis.

METHODS

Packet Visualization

Relationship to Dot Plots

Filtering

Time with animation and window duration

Selecting nodes by inclusion / exclusion

Context and Aggregation

Aggregate causality scores showed clear differences between datasets. <figure> However these views did not support well hypotheses besides that the data contained some relationships of value. We chose to allow filtering to show data at 50%, 95%, and 99% confidence levels, obtained by bootstrap sampling.

FUTURE WORK

Improved Aggregate Data Sense

We extend dot plots to paint them over where edges would be otherwise. We address the following issues: density differences for different edge lengths, overplotting due to overlap.

Performance

Precomputing totals. If multiple users analyze the same data, otherwise local computation may be hosted in the cloud. Streaming a user interface across the cloud still stutters due to latency, and prior work has shown that increased latency negatively impacts the breadth and variety of hypotheses explored during the analysis process. We can split the middle with techniques similar to those demonstrated by Immens[4], which calculates binned aggregations and sends that data for local computation and display. Additional opportunities in adapting those techniques for dynamic graph visualization would address how to

Visual Hypothesis Testing

While the current implementation supports setting up a 50/50 visual analysis test, It lacks features to share analysis tests over the network. We imagine distributing analysis keys or a unique URL to remote participants or through services such as Mechanical Turk. We could support branching trees of analysis, multiple null models. Users may also benefit from an environment to interpret the results of the tests, especially as the space of expressible test sets expands.

Visual analysis testing may produce richer results by recording the depth and other interaction characteristics that users chose from. In this work we plan to measure standard measures of analysis tasks and attempt to at best model and at least cluster the behavior patterns, to guide the interpretation of visual analysis testing results. For example, how does a person's belief in a hypothesis vary with the time spent judging it? How does this vary across individuals? What empirical distributions do the judgments of individuals follow? What if two or more people collaborate in judging a hypothesis? Some hypotheses we hope to test include: 1) time spent analyzing first increases the chance a person will judge the hypothesis true, then decreases, 2) people collaborating in person will be more likely to agree than they would compared to collaborating remotely, even when partially controlling for confounding factors such as sharing or not sharing a workstation or including high definition video conferencing, 3) individual differences in analysis styles, quantified by the amount of branching and time spent in different analysis paths, will predict their responses for certain kinds of data, and 4) casual visual hypothesis test results, aggregated over a large number of people, provide a unique signal potentially useful when combined with other, more traditional analysis methods.

We also plan to explore sonification, both in active (quick changes / small samples) and passive listening modes (long durations), as a potential way to derive a unique signal for analysis. This and other non-traditional ways for the public to interact casually with science present unique design and interaction challenges, such as along the active – passive experience dimension. Raising awareness and hopefully interest in scientific issues, especially uncertainty and validity of data, can lead to broader impacts. We are also interested in exploring other complex datasets with these techniques, such as the time-varying interactions on social networks.

CONCLUSION

ACKNOWLEDGMENTS

REFERENCES

1. Luce, Paul A, Large, Nathan. Phonotactics, density, and entropy in spoken word recognition. Language and Cognitive Processes (2001).
2. Anderson, R.E. Social impacts of computing: Codes of professional ethics. *Social Science Computing Review 10*, 2 (1992), 453-469.
3. How to Classify Works Using ACM's Computing Classification System. http://www.acm.org/class/how_to_use.html.
4. Klemmer, R.S., Thomsen, M., Phelps-Goodman, E., Lee, R. and Landay, J.A. Where do web sites come from? Capturing and interacting with design history. In *Proc. CHI 2002*, ACM Press (2002), 1-8.
5. Mather, B.D. Making up titles for conference papers. *Ext. Abstracts CHI 2000*, ACM Press (2000), 1-2.
6. Schwartz, M. *Guidelines for Bias-Free Writing*. Indiana University Press, Bloomington, IN, USA, 1995.
7. Zellweger, P.T., Bouvin, N.O., Jehøj, H., and Mackinlay, J.D. Fluid Annotations in an Open World. *Proc. Hypertext 2001*, ACM Press (2001), 9-18.

The columns on the last page should be of approximately equal length.