RESEARCH STATEMENT

Mershack Okoe

Introduction

Data visualization research draws on theory from computer graphics, human computer interaction, and cognitive science to help quantify data into interpretable and interactive visual designs. My research in data visualization focuses on two themes: building frameworks to support the design and evaluation of visualization systems, and exploring guidelines for designing effective visualizations. My overarching research goal is to support the design of tools that are effective in helping users rapidly comprehend and gain insight from their data.

I follow a user-centered approach to research with a focus on investigating ways that data visualization tools can increase the learning speed, task efficiency, and comfort level of users. I believe putting the user at the forefront during the visualization development process leads to the design of simple efficient visualizations with wide user adoption.

Building Frameworks to Support the Evaluation of Data Visualization Systems

My doctoral thesis focused on simplifying how user studies (evaluations conducted with users) are conducted in the data visualization community. User studies are widely used in the data visualization community to validate research, compare visual designs, and identify problems with visualization systems. However, the processes involved in designing and running user studies is challenging and time consuming.

In my dissertation research work, I investigated how to build a framework that improves the efficiency of designing and running web-based quantitative user studies, and how to support crowdsourced user studies. I designed the GraphUnit framework [1] which provides a simple way of designing and running graph user studies. It allows evaluators to design user studies using graph taxonomy tasks and benchmark graph datasets. It places the user studies on crowdsourcing platforms such as Amazon Mechanical Turk, manages the running of the studies, and generates statistical analyses of the study results. We demonstrated that GraphUnit can be used to create and deploy previously published graph evaluation studies in a matter of minutes.

I also designed the VisUnit framework which allows evaluators to design and run user studies of many types of visualizations. VisUnit builds on the GraphUnit framework, and allows evaluators to design user studies using their own tasks and datasets. It streamlines the user study process by simplifying the design of the study, managing the actual running of the study, and providing appropriate statistical analyses of the study results. We demonstrated the effectiveness of VisUnit by showing that it can be used to replicate the methodology of 84% of 101 controlled user studies published in IEEE Information Visualization conferences (between 1995 and 2015). These user studies involved graphs, multidimensional visualizations, trees, 2D areas, and temporal visualizations. Part of the work on VisUnit has been submitted to Eurovis 2018, and other parts are in preparation for future publications.

Future Research: I plan to continue building tools that can reduce the overhead involved in the creation and evaluation of visualization systems. There are several future research ideas that I plan to explore. First, I plan to extend VisUnit to support *beyond time and error* measurements such as measuring number of clicks, number of interactions, eye-tracking, think-aloud protocols, and user stress levels. The nature of online studies pose unique challenges to *beyond time and error* measurements and I plan to investigate novel ways to solve such challenges. Second, I plan to investigate how to standardize user studies to allow evaluators to store their user study designs and study results in publicly accessible repositories. This will provide opportunities for other researchers to easily extend existing studies with new tasks and additional variables. Thirdly, I plan to investigate the following research questions: How can we evaluate collaborative visualization systems with crowdsourcing?, and How can we streamline the design and running of longitudinal and insight-based studies?

Exploring Guidelines for Designing Effective Visualizations

Data visualization leverages the power of human vision and cognition to help users gain insight from their data. I believe the strength of data visualization lies in how best it helps users perform their intended tasks and how best it help users gain insight from their data. I follow an experimental research approach to evaluate existing and new design choices to discover visual designs and interaction techniques that take full advantage of human vision and cognition.

In our Eurovis 2014 paper [2], we worked on reducing the difficulty that users face when performing graph tasks with large and cluttered graph visualizations. We investigated how eye-tracking can be used as a real-time input to reduce clutter and support user tasks in dense graph visualizations. We showed that eye-tracking can be used to detect tasks that users intend to perform, and unobtrusive visual changes can be made to graph visualizations to support such tasks in real-time. Unobtrusive visual changes include making visual elements under a user's focus more salient than others, and filtering out visual elements irrelevant to tasks. Also in this paper, we introduced a novel fovea-based filtering algorithm that dims out unnecessary edges in a user's fovea. We also introduced techniques for correcting gaze offsets commonly reported by current eye tracking systems. Results from a controlled user study showed that users were more accurate on connectivity tasks with an eye-tracking-enabled graph visualization than they were with an eye-tracking-disabled graph visualization.

In our Graph Drawing 2017 paper [4], we presented the results of a study we conducted to compare the effectiveness of the two primary techniques for visualizing network data: Nodelinks (NL), and Adjacency Matrices (AM). We embarked on this research after finding no definitive information in the visualization literature regarding which tasks are better supported by these two techniques. We discovered from a preliminary work [3] that using ecologically valid dataset and ecologically valid interactive techniques leads to results different from results obtained by previous research. Our study involved 557 crowdsourced participants who used interactive versions of the two techniques. We found out that: NL is better than AM for questions about network topology and connectivity, NL is comparable for group and memorability tasks, and NL is a better overall choice for visualizing datasets similar in size and structure to the one we evaluated. Our work won the best paper award in the applied track at GD 2017, and we have recently submitted an extended version of our paper to Transactions on Visualization and Computer Graphics (TVCG) by invitation.

Future Research: I plan to continue exploring techniques that can inform our understanding on effective visualization design for different tasks and target audiences, and to seek ways that designers can take full advantage of human vision and cognition. I plan to continue seeking for open problems and limitations in the visualization literature, and perform task-based user studies to find techniques that can improve how novice and expert users explore and reason about their data. Another area of research I'm keenly interested in exploring is data visualization for personal usage. We now have access to immense amount of data about our communities, health and fitness, and social media networks. These data are relevant to our everyday decision making process, and I'm interested in investigating how visualization can be designed to help users rapidly understand and explore their personal data especially for users with little or no experience in data visualization.

Multi-Disciplinary Collaboration

Applications are essential in data visualization research because they help validate data visualization solutions in real-world problems. Building applications through interdisciplinary collaboration with domain experts helps in gaining background knowledge and vital feedback on domain-based designs which also influences the design of future applications. In a collaboration with Professor Angela Laird from the Neuroinformatics lab in Florida International University, we built an interactive application to help analyze data involving neuroimaging experiments mined from over 10,000 task-based brain mapping experiments. The application we developed helped our collaborator to visualize meta-analytic coactivation networks that showed how functional brain systems decompose into constituent sub-networks and how coactivation networks relate to mental functions. This work is published in NeuroImage [5].

Future work: To leverage the increasing scale and complexity of data being generated in different disciplines, there is the need for applications that help domain experts analyze their data. New techniques will need to be developed and old techniques will need to be used in novel ways. Collaboration helps drive new research directions for both parties and provides an opportunity to create real-world impact. I plan to explore opportunities for interdisciplinary collaborations on new research problems in diverse disciplines. I plan to foster collaborations with experts in biology, medicine, chemistry, psychology, and social science. It is my goal to work with these multi-disciplinary experts to design novel data visualization tools that helps them explore, generate and test hypotheses, and gain insight from their data.

References

1. Mershack Okoe, and Radu Jianu. "Graphunit: Evaluating Interactive Graph Visualizations Using Crowdsourcing". In *Computer Graphics Forum*. *Vol. 34. No. 3.* (2015).
2. Mershack Okoe, Sayeed Safayet Alam, and Radu Jianu. "A Gaze-enabled Graph Visualization to Improve Graph Reading Tasks". In *Computer Graphics Forum. Vol. 33. No. 3*. (2014).
3. Mershack Okoe, and Radu Jianu. "Ecological Validity in Quantitative User Studies–a Case Study in Graph Evaluation". In *IEEE VIS 2015 poster program.* (2015).
4. Mershack Okoe, Radu Jianu, and Stephen Kobourov. "Revisited Network Representations". In *25th Symposium on Graph Drawing*. (2017).
5. Angela R Laird, Michael C Riedel, Mershack Okoe, Radu Jianu, Kimberly L Ray, Simon B Eickhoff, Stephen M Smith, Peter T Fox, Matthew T Sutherland. "Heterogeneous fractionation profiles of meta-analytic coactivation networks". In *NeuroImage.* 2017.