Visualizing the Impact of Organizational Change

Project proposal

Rob Barwell
Carleton University
Ottawa, Canada
rob@barwell.ca

Eric Spero
Carleton University
Ottawa, Canada
eric.spero@carleton.ca

1. DESCRIPTION

We aim to design and develop a novel visualization to support managerial decision-making by helping managers understand the effects of personnel change on an organization.

2. READING REVIEW

The rise of the 'gig economy'[4, 9] has forced society through a massive transition in the past two decades with people transitioning between jobs with greater frequency. Historically, a person would obtain a job from high school or university and stay with a specific company until they retire. This resulted in organizations that had minimal change, which could be easily managed. Modern organizations are forced to adapt with the shift to the gig economy and other non-traditional employment models.

An organization is a social network[16] of individuals engaged in complex *interlocking contingencies*[10]: individuals in an organization are tightly interconnected, where the behaviour of one both depends on, and has subsequent consequences for the behaviour of others[10]. This feature of organizations means that the results of changes to its social structure (say, by adding or removing an individual) can be difficult to understand. Change is disruptive, and care must be taken to minimize any negative side-effects. The challenge of minimizing organizational disruption as a result of personnel change depends critically on understanding the effects of that change. In the gig economy, this challenge, and the need for tools that support it, is even greater.

We aim to address this problem through *information visualization*. Visualizations support thought by reducing the gap between the data, and the users' *mental model* of the data[21]. A mental model is an internal representation of how something in the world works[18, 13]. Wherever there is distance between the presentation of the data and our understanding of the data, mental work must be done so that understanding is possible. This type of mental work does not bring us closer to solving domain goals, but rather is a sort of unfortunate precursor for the really important work. This type of workload is referred to as *extraneous workload* [14]. Fortunately, the physical environment can be used to store information, which allows us to 'off-load' mental work onto the environment[20]. Visualizations are essentially one way of effectively leveraging this property of the environment to aid thought.

Visualizing organization change has many challenges. Some of these challenges are: how to represent a large organizational structure; how to provide focus on specific organizational change, without losing context of the organization; and how to allow the user to navigate the space. A novel visualization can help in many areas with this problem.

Visualization has been used to help analyze social networks since at least the 1930s[8]. Social networks are commonly depicted by point-and-line graphs such as directed graphs[8] (e.g. [2, 11]). Another method of visualizing social networks is with tree-maps[17, 15, 7]. Using a tree-map emphasizes the proportional representation of node attributes such as size[17], whereas a directed graph emphasizes the relationships. Organizational change requires an understanding of both relationships between positions and output of a specific position such as how many people they supervise. Representing nodes (people) within a visualization requires specific thought on how they are represented. One way to represent multiple attributes of a node is using glyphs[19, chapter 5]. Graph and glyphs can be used together to understanding both relationships between positions and specific information about a position. Another critical factor for visualizing relationships is the layout of an organizational structure. This requires a robust layout algorithm to ensure nodes and edges can be easily differentiated[12].

The focus+context problem is well known in information visualization. Ware presents many options for helping the user focus on specific information. These include form, color, motion, and spatial position, with the strongest effects being color, orientation, size, contrast, and motion[19, chapter 5]. An organizational chart will require the correct use of these options to not overwhelm the user. The focus+context problem can also be addressed through clustering. The advantage of clustering is allowing the user to focus on a given change by reducing the number of visible elements[12]. Clustering an organizational chart into departments allows the user to easily focus on relationships between groups and not get lost in the details. This concept can be expanded further to include dynamic filtering to remove data points which are not relevant to the user. When employing filters, they should be tightly coupled and dynamic which allows rapid, incremental and reversible changes to query parameters[1].

The addition of motion to a visualization has become commonplace and expected by most users. This aids the user in easily navigating the data to obtain the information they are seeking. An example of motion would be zoom and pan. Zooming can be further divided into geometric zooming and semantic zooming. Geometric zooming simply provides a blow up of the graph content, where semantic zooming changes the content of an area to include more detail[12].

Zoom and pan would be a preferred method compared to other options such as fisheye distortion, since users are more familiar with it. Semantic zooming complements clustering by grouping data points into departments and allowing the user to zoom in and see details if desired. Motion in 2D is easy for the user to understand, however presents challenges with occlusion in 3D. 3D visualizations were introduced with the hope that the extra dimension would provide additional space to display larger structures[12]. Given the complexities associated with occlusions in 3D, a 2D approach would be preferable for organizational charts. A novel 2D representation was presented by Becker and Cleveland called brushing scatterplots where each dimension was broken down in its own axis and plotted together on one graph[3]. The user could then navigate the scatter plot and have changes / selections within one box represented on the other boxes. This concept would help when including multiple views within the same visualization by allowing the user to select a data point in one visualization and the focus changing in another visualization.

Visual formalisms yet to discuss: Provide Overview, Adjust, Detect Pattern, Match Mental Model[21].

3. DETAILED DESCRIPTION

3.1 Domain

An organization is a complex network of individuals engaged in interlocking contingencies[10]. When the network is changed, say by adding or removing an individual, the effect this will have on the whole system is difficult to predict/understand. Change is inherently disruptive, but with careful management the negative side-effects of change can be mitigated. The better managers understand the consequences of change, the better position they are in to minimize disruption, and maximize the performance of the organization.

3.2 Tasks where visualization will help

Understanding organizational change is a large problem. This project will focus specifically on the stories that become present through a visualization and not how the underlaying data is optimized or valued.

Visualization will specifically address how the user understands the results of an organizational change.

1. What relationships are affected and how can that impact the flow of information in the process;
2. Which positions would be candidates to take on additional workload during the transition or on a more permanent basis;
3. What is the value of an employee to the organization; and
4. Other stories as they appear.

3.3 Design approach

The task of helping a user make decisions regarding organizational structure and change is complex. A well thought out visualization is key to aiding a user with their decision-making process. This includes presenting the user with both an egocentric and sociocentric view of the change. Our visualization will support this by providing three concurrent views that allow the user to navigate organizational change and support the tasks identified in the previous section. To ensure we select the best visualization for the tasks we have selected an agile development methodology as follows:

1. Select and augment base data;
2. Create a skeleton visualization for feedback;
3. Obtain informal feedback from users;
4. Refine and complete the visualization trying to accommodate feedback from the users;
5. Obtain informal feedback from users on the final draft;
6. Create the final visualization.

By following an agile development methodology, it will provide us with a challenge function to ensure we are meeting the two goals of a visualization which are: communicating information to users and discovery of new knowledge from the visualization.

The visualization will have 4 distinct panels for the user to interact with. They are number in figure 1 and named:

1. Menu;
2. Global Navigation;
3. Information Flow; and
4. Supply.

The menu section is generic and allows the user to do configuration associated with the application. This includes specifying the data set, exporting results, and providing context about which perspective the user is viewing the data from. This section is supplemental to the visualization.

The global navigation section is key for the user to search and find the data they are looking for. Organizational structures present certain challenges to visualization as Herman points highlights the "size of the graph to view is a key issue in graph visualization"[12]. To aid the user with this problem our organizational chart will include a pan and zoom function which allows the user to geometrically zoom and pan the layout. The layout will be further simplified by collapsing nodes and clustering data points together to reduce the number of nodes on the screen at once. The global navigation view is primarily used for filtering and navigation. The panel will have the ability to expand and collapse as the user focuses on other tasks and explores the data results.

To support exploratory browsing[1] additional filter criteria will be available to the user when the global navigation panel is expanded to full size. Shneiderman best described filters as rapid, incremental and reversible changes to query patterns[1]. Our visualization will allow the user to filter based on number of emails and name. These filters will instantly be seen on the visualization with the opacity change for each node. Nodes that meet criteria will be displayed at full capacity and those which do not meet criteria will be displayed at reduced opacity. This will allow the user to maintain full context while searching.

When a user hovers over a node it will display links between people in the organization which have relationships through email. Since this is a separate category of links from the traditional organizational chart relationship it will be presented in a different color as Ware emphasizes that color is "excellent for labeling and categorization"[19].

The visualization will contain many salient[19] ques for the user to determine what is the currently selected person and the history of how the user navigated the data. This includes using a separate color to show the currently selected node on the global navigation and supply panels. Additionally, links within the global navigation panel will appear in different color to denote how the user has been exploring the data. To further enhance the user experience the menu bar will include forward and backward buttons to enable quicker exploration of the data.

Each panel allows the user to interact with the visualization in a variety of ways. The main focus of the global navigation panel is allowing the user to "drill down and find more data about anything that seems important"[19]. When the user selects a node in the global navigation panel it will reorient the information flow and supply panel to the egocentric perspective of the person node selected on the global navigation view.

A tree map was chosen to depict the users flow of information on the information flow panel. This allows the data to use a space filling approach as described by Shneiderman in 1992[17]. This proportional approach allows the visualization to scale with a large organization where a person could interact with 100+ people to accomplish their job. However, people in an organization typically interact with a smaller subset more frequently. This will be visually salient[19] on the tree map with the most important relationships appearing as the largest rectangles. Color will also be used to emphasize[19] which department the people are from in the tree map to help provide an additional sociocentric perspective of which department would have the most challenges if the person was no longer with the organization. Tool tips will provide additional on demand details[6]. This will include full email, department, and could contain any other information the user is interested in.

After the user has refined the visualization to the person they are interested in and explored / understands how information flows between the person and the rest of the organization they can start to make decisions around what to do during a transition or longer period. This is supported in the visualization through the supply panel. In a typical organization the supply for these periods would most likely come from the persons own group. The bar chart was chosen to represent this as a quick way to compare the relative workload of everyone in the group. Color is used again to denote the currently selected person. The user also has the option to navigate to the perspective of another person by clicking on their bar in the chart.

The goal of our visualization is best described by Yi, Kang, Stasko, and Jacko: "harnessing human's remarkable visual perception capabilities to help identify trends, patterns, and unusual occurrences in datasets"[21]. In our specific context this is supporting the user's decision-making process about organizational change. We are confident our design meets this objective by using all the relevant elements discussed in class.

3.4 Development Technology

MySQL, PHP and D3 will be used to implement our visualization. These technologies were chosen based on their popularity for visualization development.

MySQL will be used to store and access data for the visualization. A database was required due to the volume of data and requirement to access the data in subsets based on user interaction with the visualization. The Carnegie Mellon Enron email dataset[5] will be used as the base data for the visualization. This dataset will require pre-processing as it contains the full emails of 148 users. This requires Python and Perl to extract the required information and transform it for ingest into MySQL. The data will further be augmented with additional attributes that are generated for illustrative purposes to support the visualization. Additional data processing will be limited since the class is focused on visualization of data and not data modeling. Additionally, the value model chosen will be limited to number of emails sent and not a more elaborate model requiring additional data sets and data manipulation.

PHP is required to translate the data from a relational database model into JSON. JSON is the primary input format used by the D3 JavaScript library. The D3 visualization library was chosen for many reasons. It is a robust library offering many different visualization options along with rich documentation for implementing it. It was originally released in 2011 and has become a stable library.

Eric has a moderate understanding of D3 and increasing his knowledge of MySQL and PHP. Rob has used MySQL, Perl, Python and PHP extensively in the past and increasing his knowledge of D3 rapidly.

3.5 Project Plan

26 Oct 2018	Submission of Project Proposal
04 Nov 2018	Complete Preliminary Visualization
11 Nov 2018	Complete Review of Visualization
18 Nov 2018	Complete Refinements Based on Visualization Review
25 Nov 2018	Complete 2nd Review of Visualization
30 Nov 2018	Complete Refinements and Presentation
14 Dec 2018	Submit Final Paper

References

[1] Christopher Ahlberg and Ben Shneiderman. "Visual information seeking: Tight coupling of dynamic query filters with starfield displays". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 1994, pp. 313–317.

[2] Mathieu Bastian, Sebastien Heymann, Mathieu Jacomy, et al. "Gephi: an open source software for exploring and manipulating networks." In: *Icwsm* 8.2009 (2009), pp. 361–362.

[3] Richard A Becker and William S Cleveland. "Brushing scatterplots". In: *Technometrics* 29.2 (1987), pp. 127–142.

[4] Valerio De Stefano. "The rise of the just-in-time workforce: On-demand work, crowdwork, and labor protection in the gig-economy". In: *Comp. Lab. L. & Pol'y J.* 37 (2015), p. 471.

[5] *Enron Email Dataset*. 2015. URL: `https://www.cs.cmu.edu/~./enron/` (visited on 10/22/2018).

[6] Ana Figueiras. "Towards the Understanding of Interaction in Information Visualization". In: *2015 19th International Conference on Information Visualization* (2015), pp. 140–147.

[7] Terrill Frantz and Kathleen Carley. "Treemaps as a tool for social network analysis". In: (2005).

[8] Linton C Freeman. "Social network visualization, Methods of". In: *Computational Complexity*. Springer, 2012, pp. 2981–2998.

[9] Gerald Friedman. "Workers without employers: shadow corporations and the rise of the gig economy". In: *Review of Keynesian Economics* 2.2 (2014), pp. 171–188.

[10] Sigrid S Glenn and Maria E Malott. "Complexity and selection: Implications for organizational change". In: *Behavior and Social Issues* 13.2 (2006), pp. 89–106.

[11] Derek L Hansen, Ben Shneiderman, and Marc A Smith. *Analyzing social media networks with NodeXL: Insights from a connected world*. Morgan Kaufmann, 2010.

[12] Ivan Herman, Guy Melançon, and M Scott Marshall. "Graph visualization and navigation in information visualization: A survey". In: *IEEE Transactions on visualization and computer graphics* 6.1 (2000), pp. 24–43.

[13] Donald A Norman. "Some observations on mental models". In: *Mental models*. Psychology Press, 2014, pp. 15–22.

[14] Fred Paas, Alexander Renkl, and John Sweller. "Cognitive load theory and instructional design: Recent developments". In: *Educational psychologist* 38.1 (2003), pp. 1–4.

[15] Mithileysh Sathiyanarayanan and Nikolay Burlutskiy. "Visualizing social networks using a treemap overlaid with a graph". In: *Procedia Computer Science* 58 (2015), pp. 113–120.

[16] John Scott. "Social network analysis". In: *Sociology* 22.1 (1988), pp. 109–127.

[17] Ben Shneiderman. "Tree visualization with tree-maps: 2-d space-filling approach". In: *ACM Transactions on graphics (TOG)* 11.1 (1992), pp. 92–99.

[18] Nancy Staggers and A. F. Norcio. "Mental models: concepts for human-computer interaction research". English. In: *International Journal of Man-Machine Studies* 38.4 (1993), pp. 587–605.

[19] Colin Ware. *Information visualization: perception for design*. Elsevier, 2012.

[20] Margaret Wilson. "Six views of embodied cognition". In: *Psychonomic bulletin & review* 9.4 (2002), pp. 625–636.

[21] Ji Soo Yi et al. "Toward a deeper understanding of the role of interaction in information visualization". In: *IEEE Transactions on Visualization & Computer Graphics* 6 (2007).