Visualizing the Impact of Organizational Change

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

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

October 28, 2018

1 Description

We plan 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'[5, 10] 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[17] of individuals engaged in complex *interlocking contingencies*[11]: individuals in an organization are tightly interconnected, where the behaviour of one both depends on, and has subsequent consequences for the behaviour of others[11]. 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[22]. A mental model is an internal representation of how something in the world works[19, 14]. 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, and therefore should be avoided wherever possible[15]. Fortunately, the physical environment can be used to store information, which allows us to 'off-load' mental work onto the environment[21]. 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 micro-level organizational changes without losing context of the organization; and how to allow the user to navigate the large information space. As with all graphs, a robust layout algorithm must be selected to ensure nodes and edges can be easily differentiated[13].

Visualizations have been employed to help analyze social networks since the 1930s[9]. Social networks are commonly depicted by point-and-line graphs such as directed graphs[9] (e.g. [2, 12]), and also with tree-maps[18, 16, 8]. Using a tree-map emphasizes the proportional representation of node attributes such as size[18], whereas a directed graph emphasizes the relationships between nodes. Understanding the effects of organizational change requires an understanding of both relationships between individuals, and the productive output a specific individual, such as how many people they supervise. In the terminology of the domain of social network analysis[4], a visualization solution to the problem of understanding organizational change will require both *egocentric* and *sociocentric* representations of social networks. An egocentric social network is a network centered around a single individual, whereas a sociocentric network focuses on the larger social system that embeds individuals.

Individuals in an organization has multiple attributes which must be considered by the user. One way to represent multiple attributes of a node is using glyphs[20, chapter 5]. Graph and glyphs can be used together so the visualization represents both relationships between positions, and the productive output of a position.

The 'focus+context' problem is well known in information visualization. Ware[20] presents many options for helping the user focus on specific information. These include form, colour, motion, and spatial position, with the strongest effects being colour, orientation, size, contrast, and motion[20, 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[13]. 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 for navigational purposes is a common strategy for making the visualization of large amounts of data manageable. Zoom and pan is one particularly successful strategy. 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[13]. 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.

Most visualizations are in 2D. 3D visualizations were introduced with the hope that the extra dimension would provide additional space to display larger structures[13]. However, 3D visualizations introduce an occlusion problem, where nodes can become 'hidden' behind other nodes, which harms the aesthetics and usability of the visualization[13]. Given this disadvantage, a 2D approach would be preferable for visualizing social networks in an organizations.

The use of multiple views in the user interface has been used to give the user multiple simultaneous vantage points of a dataset. In Becker and Cleveland's *brushing scatterplots*[3], for example, each dimension was broken down in its own axis and plotted together on one graph. 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.

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”[22]. In our specific context this is supporting the user’s decision-making process about organizational change.

3 Detailed Description

3.1 Domain

An organization is a complex network of individuals engaged in interlocking contingencies[11]. 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

Managers require many tools to help better understand change within their organization. This requires looking at the problem from multiple perspectives prior to enacting any change to ensure it does not have unintended consequences. Prior to presenting a visualization to the user the system would pre-process data and add additional information such as a value model. The pre-processing tasks are beyond the scope of this project and we will focus only on tasks where visualization can help.

Process analysis Organizations are built upon many inter-relating processes. Within these processes relationships between people are key. These relationships enable large processes to be accomplished by splitting them into smaller tasks which are manageable by a single person. However, when a person is removed from the process it could cause the process to break. This would require the manager to find a solution to continue the process. The employees attached to a given process are typically tracked through documentation, but in practice this documentation is often lacking and out of date, in part due to the high turnover rate. A novel visualization such as the one we propose could help identify the inter-connecting relationships from a given person within an organization to determine which processes might fail if they were removed.

Allocation of human resources After identifying the inter-connecting relationships from a given person in the organization, a manager can look at options to fill the gaps that remain. Typically, this includes re-distributing the workload over the remaining individuals within a section of the organization. Therefore, the visualization would need to include a method to show options for people to fill the gaps that were identified when the person was removed from the organization.

Assessing value to organization The previous two tasks focus on the role a single individual has on a process—i.e. from an egocentric viewpoint. This viewpoint is helpful in determining what gaps need to be filled when a particular individual is removed from a process. However, such a viewpoint is not well-suited for communicating the value an individual has to the organization. In cases like this, a more holistic (or sociocentric) visualization could be helpful to managers as well.

Miscellaneous / potential future work Our group discussed other tasks that would be helped by visualizations, however we decided to focus on the three listed above. Future papers could cover additional stories such as redundancy identification by identifying people who have similar overlaps in relationships and interactions. Additional research could be done in defining value models for employees to organizations and how to visually present the worth of an employee to help with retention options. We also expect during our development to identify new use cases for our visualization.

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 four distinct panels, each allowing the user to interact with the visualization in a different way. Figure 1 shows a screenshot of the interface with the four panels labelled. They are:

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

3.3.1 Menu panel

The *menu panel* is a generic menu, and allows the user to configure the application such as 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 right section of the menu bar will be dedicated to providing the user with a capability to quickly reverse their navigation path. This ability is key to enabling the user's seamless exploration of the data set. The user will navigate the dataset by selecting a person in the organization to change the perspective of the visualization to. This change in perspective will add the current person to a dropdown list in the menu bar. This list will contain the last 25 people the user has selected. The forward and backward buttons beside the

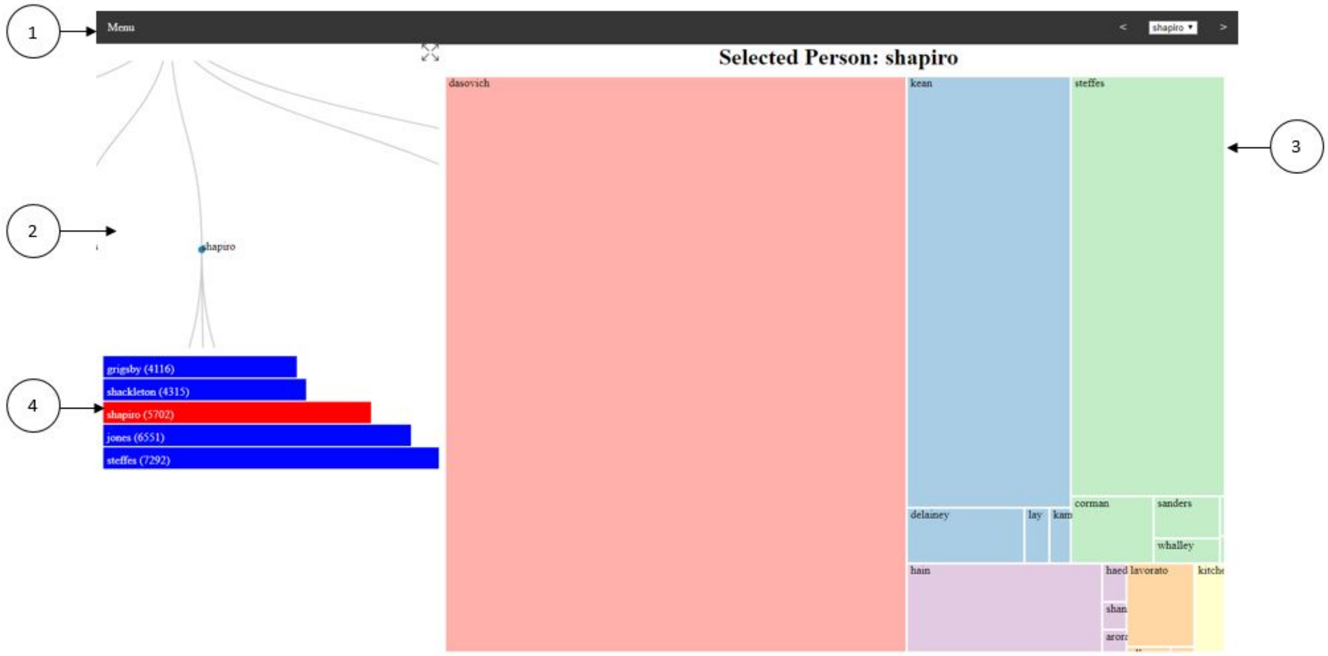

Figure 1: Four panels: 1. Menu; 2. Global navigation; 3. Information flow; 4. Supply

list will allow the user to quickly navigate to previous steps in the users' discovery process. The user will also have the option to select a person from the drop down, which will quickly re-orient the visualization perspective to the person selected from the drop down by the user.

3.3.2 Global navigation panel

The *global navigation panel* presents a tree graph linking people together in an organizational hierarchy. 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. This panel allows users to "drill down and find more data about anything that seems important"[20]. Organizational structures are typically quite large, so we must keep the size of the visualization manageable—a classic problem in tree graphs[13]. To aid the user with this problem, users will be able to pan and zoom the organizational chart to hone in on areas of particular interest as shown in Figure 2. Additionally, we will cluster data together, and nodes in the tree will be collapsible and expandable. Only the first three levels of the organization chart will be displayed when it is initially rendered. Nodes will remain in the expanded or collapsed state unless the user explicitly changes it. Collapsed nodes in the global navigation panel communicate limited information to the user. To provide additional information each node will be proportionally scaled based on the sum of all emails generated by the node and its children. This will provide additional information such as the relative comparison between the value of one organizational unit versus another. The global navigation panel will also be expandable and collapsible on demand. This allows the user to quickly switch tasks between navigating and discovering people and analyzing and evaluating the results without losing valuable screen real estate to either process.

Further support for the problems associated with displaying large amounts of data will be provided in the form of filter criteria. Our visualization will allow the user to filter based on two criteria: number of emails and name. A more fulsome real world data set would contain all the attributes associated with a node which are stored in the central address book of an email system. The results of the filter will be

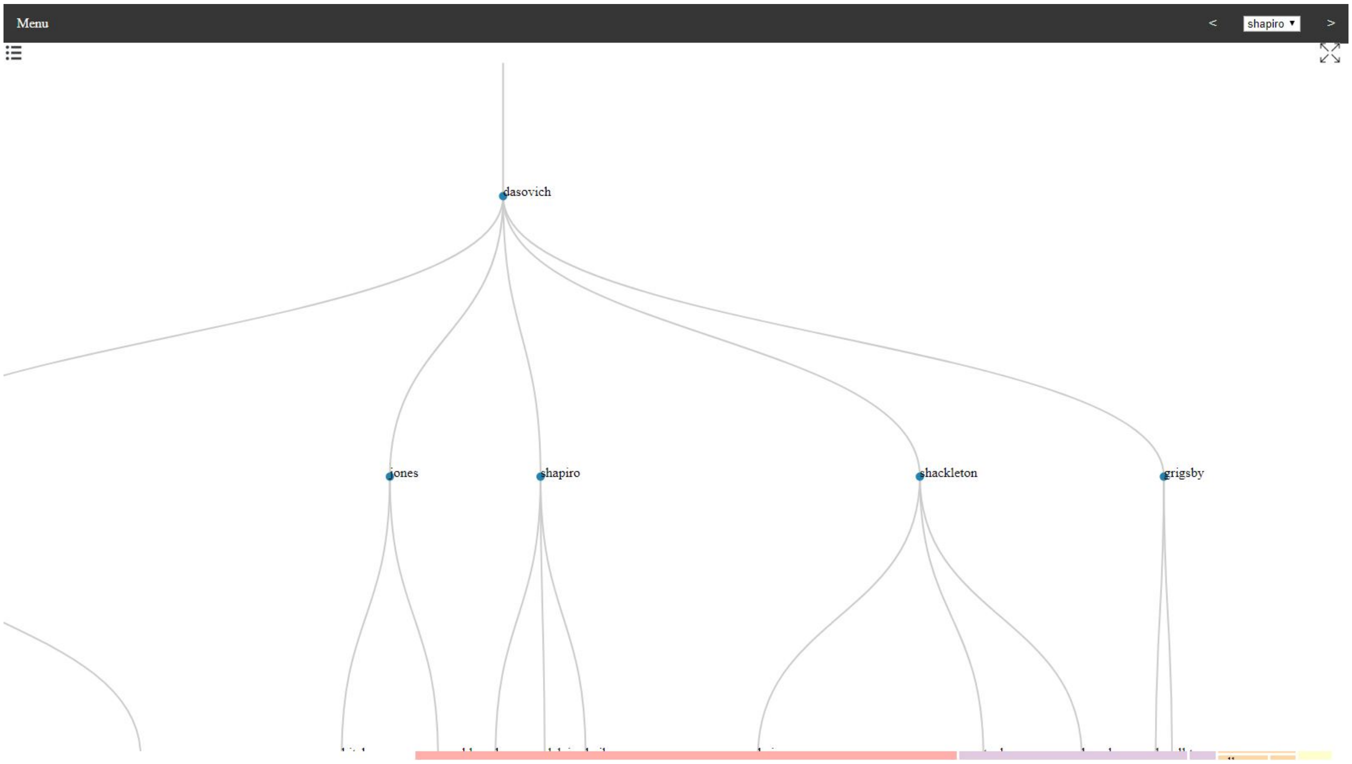

Figure 2: Global navigation panel, expanded

applied instantly, where nodes not matching the filtering criteria will be de-emphasized by a reduction in opacity. We chose to de-emphasize nodes instead of removing them entirely to preserve context. Consistent with Shneiderman's rules for filters, our filters will be rapid, incremental and reversible changes to query patterns[1].

When a user hovers over a node in the organization chart it will display links between people in the organization who have formed relationships through email. The relationships uncovered by analysis of email communication is of a different type than the relationships expressed in an organizational chart. Colour is "excellent for labeling and categorization"[20], so we communicate to the user that the two relationships are different by colouring the links differently in each case.

We use colour to saliently cue[20] information in the global navigation panel in at least two other ways: by colouring the selected node differently from the non-selected nodes, and by using colour to trace the user's path through the nodes in the global navigation panel.

3.3.3 Information flow panel

The *information flow panel* contains a treemap of individuals with whom the selected person has a relationship. We chose a tree map to depict the relationships between employees for two reasons. First, its proportional nature allows the visualization to scale from a very small organization, to a very large organization where a person could interact with 100+ people on a regular basis. Second, it communicates the strength of relationship in a salient manner through the size of the rectangles. The treemap will also be interactive and allow the user to drill down into each department they interact with to refocus the visualization on a subset of the relationships. This will highlight to the user people with smaller relationships that may be obscured on the higher level treemap.

We will use colour to group together individuals from the same department to help users easily understand which departments would be most affected if that person was no longer with the organization. Additional details on people[7] such as email address and department name will be provided through tool tips.

An alternative use of color for the treemap was discussed within our group to depict hierarchy instead of department. This would highlight those relationships which the organization considers more important vs which department the person interacts with. Both these pieces of information are important to a manager when they are going through their decision making process. This alternative view could be implemented using a drop down to allow the user to switch between views, or we could pick whichever is more salient through development and trials.

3.3.4 Supply panel

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 begin the managerial decision-making process. The *supply panel* presents all the users in the selected user's group. We present the people in the selected user's group because replacements for a given employee will typically come from the same department of that employee. Each user is presented as a bar in a bar chart, where the height of the bar is determined by the total number of emails sent by that user (which is an indication of their relative workload). Since a large number of people might be in a given department the bar chart will center on the selected person with the ability to scroll up and down to see everyone. Again, we use colour to denote the selected person. By clicking on the bar chart, the user navigates to that user, and the other panels update accordingly.

3.4 Development Technology

Our visualization could be implemented using any number of technologies, however we have chosen a JavaScript based stack. This choice allows the visualization to be cross platform and take advantage of client-side computer resources such as a GPU. A server will be required to transform the stored data into a form that is digestible by front end libraries.

Our final visualization will require the use of: MySQL, PHP, D3, Perl, and Python.

Pre-processing of project data is required prior to ingest and storage in a database. The Carnegie Mellon Enron email dataset[6] contains the email boxes of 148 users. This real-world data set will be the basis for our project to demonstrate the power of our visualization. During the pre-processing phase Python and Perl are required for parsing the raw email format and extracting the SMTP header information we require for our project. The SMTP header information will be augmented with additional attributes that are required to illustrate the various features of our visualization. The processed data will be stored in a MySQL database. A database was required due to the volume of data and a requirement to access the data in subsets based on user interaction with the visualization. Additional data preprocessing will be limited since the class is focused on visualization of data and not data modeling. 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 will be used to transform the data model stored in the MySQL database to JSON, which is the primary input format used by D3. PHP will be further leveraged to do any server-side dynamic web development.

We conducted a thorough review of JavaScript visualization frameworks. Our research concluded that D3 was the most widely used and mature framework. After reviewing D3 it met many criteria that we were searching for in a visualization framework. This included a mature framework that would allow us to focus on implementing our visualization without fear of running into bugs or incomplete code. It provides the ability to customize every detail of a visualization and its interaction or use high level abstractions that produce generic visualizations out of the box. Lastly, it provided all the visualizations we required in one library. We require hierarchical tree view, tree maps, and bar charts. D3 visualizations come default with a number of features such as zoom and pan, and built in methods to access click functions. Together this package represents a rapid way to implement our visualization in a robust package that has been tested to work across many platforms.

Eric has a moderate understanding of D3 and increasing his knowledge of MySQL and PHP. Rob has extensive experience with MySQL, Perl, Python, PHP 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

Two main work tasks have been completed. The data has been pre-processed to ensure we had all the required elements for our visualization. And a skeleton of the visualization has been implemented to verify the D3 library would meet our requirements. This will serve as a basis for the continued development of our visualization.

During the development of the project proposal Eric focused on preparing the written document and Rob focused on the technological prototype. Overall it was a collaborative effort between the team with each member reviewing and contributing to all the work. For example the skeleton visualization has been through a couple iterations based on feedback and refinements by Eric. Conversely Rob has reviewed and contributed to sections of the project proposal such as design and development technologies.

References

[1] Christopher Ahlberg and Ben Shneiderman. 1994. 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, 313–317.

[2] Mathieu Bastian, Sebastien Heymann, Mathieu Jacomy, et al. 2009. Gephi: an open source software for exploring and manipulating networks. *Icwsm*, 8, 2009, 361–362.

[3] Richard A Becker and William S Cleveland. 1987. Brushing scatterplots. *Technometrics*, 29, 2, 127–142.

[4] Kenneth KS Chung, Liaquat Hossain, and Joseph Davis. 2005. Exploring sociocentric and egocentric approaches for social network analysis. In *Proceedings of the 2nd international conference on knowledge management in Asia Pacific*, 1–8.

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

[6] 2015. Enron email dataset. Retrieved 10/22/2018 from `https://www.cs.cmu.edu/~./enron/`.

[7] Ana Figueiras. 2015. Towards the understanding of interaction in information visualization. *2015 19th International Conference on Information Visualization*, 140–147.

[8] Terrill Frantz and Kathleen Carley. 2005. Treemaps as a tool for social network analysis.

[9] Linton C Freeman. 2012. Social network visualization, methods of. In *Computational Complexity*. Springer, 2981–2998.

[10] Gerald Friedman. 2014. Workers without employers: shadow corporations and the rise of the gig economy. *Review of Keynesian Economics*, 2, 2, 171–188.

[11] Sigrid S Glenn and Maria E Malott. 2006. Complexity and selection: implications for organizational change. *Behavior and Social Issues*, 13, 2, 89–106.

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

[13] Ivan Herman, Guy Melançon, and M Scott Marshall. 2000. Graph visualization and navigation in information visualization: a survey. *IEEE Transactions on visualization and computer graphics*, 6, 1, 24–43.

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

[15] Fred Paas, Alexander Renkl, and John Sweller. 2003. Cognitive load theory and instructional design: recent developments. *Educational psychologist*, 38, 1, 1–4.

[16] Mithileysh Sathiyanarayanan and Nikolay Burlutskiy. 2015. Visualizing social networks using a treemap overlaid with a graph. *Procedia Computer Science*, 58, 113–120.

[17] John Scott. 1988. Social network analysis. *Sociology*, 22, 1, 109–127.

[18] Ben Shneiderman. 1992. Tree visualization with tree-maps: 2-d space-filling approach. *ACM Transactions on graphics (TOG)*, 11, 1, 92–99.

[19] Nancy Staggers and A. F. Norcio. 1993. Mental models: concepts for human-computer interaction research. English. *International Journal of Man-Machine Studies*, 38, 4, 587–605.

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

[21] Margaret Wilson. 2002. Six views of embodied cognition. *Psychonomic bulletin & review*, 9, 4, 625–636.

[22] Ji Soo Yi, Youn ah Kang, John T Stasko, Julie A Jacko, et al. 2007. Toward a deeper understanding of the role of interaction in information visualization. *IEEE Transactions on Visualization & Computer Graphics*, 6.