Visualizing multi-channel networks

Paweł Antemijczuk	Marta Magiera	Sune Lehmann	Andrea Cuttone	Jakob Eg Larsen
Technical University of Denmark	Technical University of Denmark	Technical University of Denmark	Technical University of Denmark	Technical University of Denmark
DTU Compute Matematiktorvet 303B	DTU Compute Matematiktorvet 303B	DTU Compute Matematiktorvet 303B	DTU Compute Matematiktorvet 303B	DTU Compute Matematiktorvet 303B
Denmark 2800, Kgs Lyngby,	Denmark 2800, Kgs Lyngby,	Denmark 2800, Kgs Lyngby,	Denmark 2800, Kgs Lyngby,	Denmark 2800, Kgs Lyngby,
pantemijczuk @gmail.com	marmagiera @gmail.com	sljo@dtu.dk	jaeg@dtu.dk	ancu@dtu.dk

ABSTRACT

In this paper, we propose a visualization to illustrate social interactions, built from multiple distinct channels of communication. The visualization displays a summary of dense personal information in a compact graphical notation. The starting point is an abstract drawing of a spider's web. Below, we describe the meaning of each data dimension along with the background and motivation for their inclusion. Finally, we present feedback provided by the users (31 individuals) of the visualization.

Keywords

Multi-channel data, social networks, visualization, personal data, social interaction.

1. INTRODUCTION

Visualizing social connections is a recurring subject in the field of network science. Researchers like to view them in order to get a general overview of the network, before performing in-depth analysis, but even regular people like to view their own social networks, sometimes learning something about themselves, that they did not previously realize [1].

Visualizations most often take form of graphs with nodes denoting people involved in the network and edges showing the established social connections. This form is constantly refined and attempts are made to make it more clear and readable, especially for larger networks [2]. But simply visualizing nodes and connections between the users is not everything. Both people and their relationship, so both nodes and edges can have certain attributes. By drawing an edge, only one information is conveyed – that the two people know each other (unless it is a directed graph, in which one person may claim to know the other, but not vice versa). However a social connection is much more than a binary fact and by simplifying it as such, there is a loss of information

[3]. The connection can carry a wealth of data, such as channel or mode of communication (face-to-face, via a phone call, through a social networking site), the frequency of contacts as well as such things as geographical location of contact or the mood of the conversation. This information is potentially of high interest to both scientists and the participating subject.

In our approach to visualizing social connections, we would like to focus on smaller sub-networks, but attempt to visualize as many attributes of the connections as possible, while placing less emphasis on the node attributes. We hope that this will provide a unique insight on the user's own connections and provide interesting and stimulating self-feedback.

2. RELATED WORK

There are many tools that create visualizations of networks, such as have been mentioned before. Most of them focus on large, sprawling networks and attempt to display them in a clear way. ContactMap [4] is an example of one of the earliest attempts of sorting an individual's social contacts in a more concise, clear and organized way. However the majority of visualizations focus only on showing the structure of the network. While some tools attempt to encode various information using attributes such as color, positioning and shape of the drawn nodes [5][6] this information usually pertains to the structure of the network, such as the community they belong to [4], the amount of connections (the degree of the node) or network distance [5].

Alternatively, certain static information can be included, such as gender, organization the person belongs to, their city of residence, etc. Displaying all that information is difficult with the limited transformations that can be applied to a node, so other system utilize panes to display different node attributes or draw expandable overlapping nodes for each of the node's attributes [7].

This still does not allow an individual to view attributes of the edges (connections), which Schneiderman et al. define as one of the six main challenges of network visualization [8].

Not many of these approaches focus on "multi-layered" social networks, (which is the case not only in visualization but also general network analysis) [9]. In such networks, different forms of contacts form different connections. The common approach is to construct several networks for each of the attributes and then combining them [3]. In principle this creates a network, which edges of have several attributes (or lack of thereof). The analytical approach does not make visualization any simpler and, in fact, combining several networks with use of color-coded edges to denote different types of links tends to be cluttered. Thus most tools fail to convey more attributes of both nodes and edges [7].

Solving the challenge, however, becomes easier as the size of the network decreases, especially if we are mostly interested in displaying detailed information about a single individual in the network (the subject of the visualization).

Additionally there is an increasing interest in self-quantified data and visualization that would aid in the process of learning about oneself [1] [10].

3. THE VISUALIZATION

In our visualization we focus on subsets (16 contacts) of an individual's social network, having built the said network out of complete information about the said individual – that is knowing about each and every social contact occurrence during a set period of time, its type (e.g. a phone call or a face-to-face meeting), duration, location and a detailed timestamp. We attempt to convey the most information possible about the individual's connections, narrowing them down to the most frequent ones or people that he or she had spent the most time with (as detailed later). Three data channels can be clearly seen in the dataset – contacts being made using the Bluetooth probe (implying face-to-face meeting), text messages as well as phone calls. Each of the three forms a separate social networks of contacts for the given user.

Instead of drawing a traditional graph we visualize the contacts in a different manner – as a metaphor of a spider's web. We would represent the user as the

"spider" sitting in the middle of the web and social contacts would be "caught in the web" in various locations of it, depending on certain parameters.

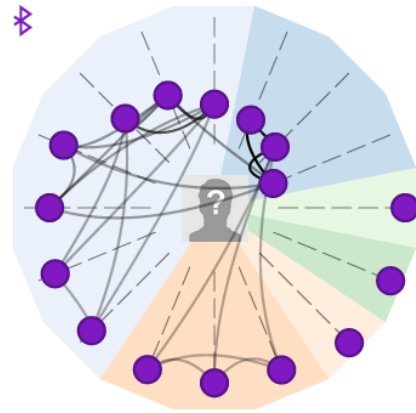Figure 1 shows the final webchart. The chart displays only one channel.

Figure 1. The final webchart displaying 16 friends of a user with which he made contacts using Bluetooth. The friends are linked based on their knowledge of each other and are grouped into communities.

As previously, each circle represents a single person the user has made contact with using the channel that the chart represents. In this case, they are Bluetooth contacts. In the middle of the chart there is a picture (or an avatar) of the user – for the testing purposes we used a placeholder picture displayed when the user does not have a picture in the system. The closer a circle is to the picture in the middle, the more contact has been made with the given person. The radial axes are meant to show the range of the values of the amount of contacts the user made with all his friends. The scale is between the minimal and maximal values for contact amounts between all friends within that channel

The webchart is built with use of radial axes and links between the circles. The links represent the connections between the friends of the user themselves using the same channel. To draw these connections, first we construct a regular graph for all social contacts in the network excluding the user. After the graph has been created using all data, less significant links are removed using a thresholding algorithm described by Serrano et al. [11] This ensures we only show the most significant links.

After the network has been made and links pruned, community detection is ran on it, using our implementation of the Louvain method [12] based on the Python implementation in the NetworkX package[1]. The background color of the axes is chosen based on the community the given friends belong to.

---

[1] http://networkx.github.io/

3.1.1 Small-multiple

The biggest challenge that we faced throughout the various iterations of the visualizations was how to display all three data channels for the user's social contacts. The data cannot be directly compared as each of the communication forms are different, however we would still like to show them simultaneously as well as allow people to make indirect comparisons between the contacts based on all three channels. In order to facilitate this we have decide to use the "small-multiple" concept as introduced by Tufte [13]. Instead of drawing one web, we would draw a web chart for each of channels used, as shown in Figure 2.

Figure 2. Small-multiple webcharts each for a different channel in the network – sms, Bluetooth contacts and phone calls.

This allows the user to compare their contacts across the channels, without introducing direct comparison between that data itself. This is possible as each contact remains on the same axis on each of the charts. In the figure we have concealed the community information for clarity. Note that the other channels do not have contact information at all; due to privacy reasons we are unable to show text messages and phone calls between the friends of the user. For Bluetooth, in fact, only contacts made in the presence of the user are recorded.

In case no data is not present for a channel, then the circle is not present on that chart.

3.1.2 Timeline

Below the multiple web charts we have placed a timeline that allows the user to change the period of time that the data is read for. If the period is changed, the circles slide across the axes to their new positions according to new data. Each chart shows the top 16 contacts (according to the sum of all contacts) for the given period.
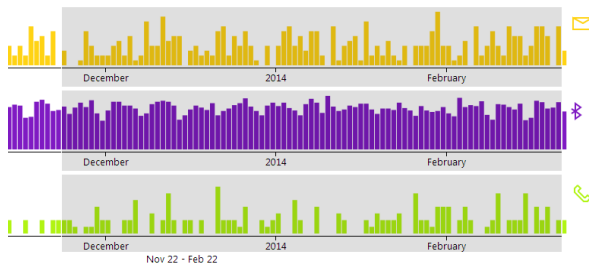

Figure 3. The timeline with barcharts for each channel showing total contacts made using the

given channel. The grey area signifies the time period chosen to be displayed on the webcharts.

The timeline is built out of bar charts – one for each channel. The bar charts show the total amount of contacts for the given; each bar being a single day. This provides a good overview of each channel's usage over time and allows us better to answer the main question about the user's data – "How does the user use his channels?".

4. EVALUATION

We have created test data sets in order to evaluate the visualization. The test data set that we tested against contained 16 participants grouped into 4 communities of friends that frequently meet during the weekday as well as 2 communities that represent friends meeting during the weekend. The communities slightly overlap. Additionally we have chosen a number of users to have much higher rate of contact using each channel.

After loading the data in the visualization the communities were detected correctly, as well as top contacts. We could clearly see that the test groups we have created in our data generators had the same background color on the visualization, that signified the communities. The contacts we have given the highest probabilities in the test generator, surfaced as closest to the center of the visualization.

4.1.1 User feedback

We have distributed the visualization along with an online survey to 31 people out of our friends and acquaintances. We asked them to evaluate the visualization, identify its features and the information that is conveyed.

In the first set we asked the users to rate on a scale from 1 to 5 (where 1 is the worst, 3 neutral and 5 the best) whether the visualization is: clear and understandable, easy to use, fun, novel and working as expected. The users responded mostly positively on all those questions with 74.2% finding it clear and understandable, 83.87% finding it easy to use, 58.06% fun, 80.64% novel and 77.42% working as expected. 9.68% of respondents found it completely unclear.

In the second section we asked the users to identify what information is conveyed by the visualization. We asked about several things, as well as some information that is **not** conveyed by the visualization in order to identify correct answers. The breakdown is as follows:

- 55% were able to identify their best friends.
- Almost 95% identified which of the contacts is called the most.
- 82% correctly identified in which days the user texts the most

- 63% correctly identified which contacts know each other
- 53% correctly identified which friends are good friends with each other (implying the community)
- 66% correctly identified what fraction of total contacts in given day, some chosen contacts are

Interestingly, even though no direct comparison are made, 79% of respondents identified which channel is used the most. This can be inferred from the popup information displaying the details, as well as the bar chart being far more uniform (contacts are made every day using that channel, not so much others). This allows for certain comparisons being made, without any direct comparisons.

There was a number of people who identified incorrect information, such as location of meetings, which implies they did not understand the visualization at all.

Lastly we asked them for any comments. In general the feedback was positive. Additionally other small suggestions were made regarding colors used and small clarifications.

5. CONCLUSIONS

We have successfully created a visualization of social contacts that is able to display three communication channels at once. It allows the user to display three (although there is nothing that would prevent from this model being used for more channels) different layers of their social network including detailed information about them.

The feedback was largely positive, with the only remarks being about the discovery of some features. This implies certain cosmetic changes might be necessary to make some features (especially the timeline changes) easier to discover.

This model can be successfully used to display multi-channel social networks, while our data contains only three channel and full interactions for only one of them, it is easily adapted for much more.

REFERENCES

[1] I. Li, Y. Medynskiy, J. Froehlich and J. E. Larsen, "Personal informatics in practice: improving quality of life through data," *CHI '12 Extended Abstracts on Human Factors in Computing Systems, CHI EA '12, 2799,* pp. 2799-2802, 2012.

[2] C. Dunne and B. Shneiderman, "Motif Simplification Improving network Visualization Readability with Fan and Parallel Glyphs," Human-Computer Interaction Lab Tech Report HCIL-2012-11, University of Maryland, 2012.

[3] B. T. Dai, F. C. T. Chua and E. Lim, "Structural Analysis in Multi-Relational Social Networks," in *SIAM International Conference on Data Mining (SDM2012)*, 2012.

[4] B. A. Nard, S. Whittaker, E. Isaacs, M. Creech, J. Johnson and J. Hainsworth, "Integrating Communication and Information Through ContactMap," *Communications of the ACM vol. 45 issue 4,* pp. 89-95, April 2002.

[5] J. Heer and D. Boyd, "Vizster : Visualizing Online Social Networks".*Proceedings of the Proceedings of the 2005 IEEE Symposium on Information Visualization (INFOVIS '05). IEEE Computer Society, Washington, DC, USA, 5-. DOI=10.1109/INFOVIS.2005.39 http://dx.doi.org/10.1109/INFOVIS.2005.39* .

[6] U. Brandes and D. Wagner, "Analysis and Visualization of Social Networks," 2001.

[7] A. Bezerianos, F. Chevalier, P. Dragicevic, N. Elmqvist and J. Fekete, "GraphDice: A System for Exploring Multivariate Social Networks," *Computer Graphics Forum, 29 doi: 10.1111/j.1467-8659.2009.01687.x,* p. 863–872, 12 August 2010.

[8] B. Schneiderman and A. Aleks, "Network visualization by semantic substrates," *Visualization and Computer Graphics, IEEE Transactions on 12.5,* pp. 733-740, 2006.

[9] P. Kazienko, P. Brodka, K. Musial and J. Gaworecki, "Multi-Layered Social Network Creation Based on Bibliographic Data," *Social Computing (SocialCom), 2010 IEEE Second International Conference doi: 10.1109/SocialCom.2010.65,* pp. 407,412, 20-22 August 2010.

[10] J. Brophy-Warren, "The New Examined Life.," *The Wall Street Journal,* 6 December 2008.

[11] M. Á. Serrano, M. Boguñáb and A. Vespignanic, "Extracting the multiscale backbone of complex weighted networks," *PNAS vol. 106,* pp. 6483-6488, 21 April 2009.

[12] P. D. Meo, E. Ferrarax, G. Fiumara and A. Provetti, "Generalized Louvain method for community detection in large networks," *CoRR, abs/1108.1502,* 2011.

[13] E. R. Tufte, The Visual Display of Quantitative Information. Second edition, Cheshire, Connecticut: Graphics Press.