

A Multiple-Aspects Visualization Tool for Exploring Social Networks

Jie Gao, Kazuo Misue, and Jiro Tanaka

Department of Computer Science, Graduate School of System and Information Engineering,
University of Tsukuba
{gao,misue,jiro}@iplab.cs.tsukuba.ac.jp

Abstract. Social network analysis (SNA) has been used to study the relationships between actors in social networks, revealing their features and patterns. In most cases, nodes and edges in graph theory are used to represent actors and relationships, and graph representations are used to visually analyze social networks. However, many visualization tools using network diagrams tend to depict most information about social networks by using the properties of nodes, which result in a visual burden when identifying actors or relationships according to certain properties. There is a lack of tools to support work by investigators to provide insights into multiple-aspect networks. We considered actors, relationships, and communities to be three important elements, and developed a tool called *MixVis* that integrates a tagcloud, network diagrams, and a list to show the elements. Our tool allows users to explore social networks from elements of interest, and acquire details through links with the three different viewpoints.

Keywords: Social network analysis, visualization, human interface.

1 Introduction

Social network analysis (SNA) is used to investigate the interactions between one person and another, one company and another, or one country and another. By using SNA, investigators attempt to find an important actor (person, company, or country), to understand the information flow and to master the trends in the whole network. SNA can facilitate an investigator's decision-making processes. SNA has been applied to many areas such as social science, politics, and psychology.

Actors, links, and communities are the main factors in social networks. A community can be defined in two different ways. In some research, a community is made up of an exclusive disjoint group of actors, and often extracted by using a clustering algorithm such as Newman's method [6]. However, in the real world, the word community can also mean an actor's affiliations. The difference between the two definitions is that a person can only join in one exclusive group, but can have multiple affiliations. The latter has recently been attracting a great deal of attention with the increasing popularity of the Social Networking Service (SNS), where end users can interact with friends by participating in a variety of virtual communities.

Nodes and edges in graph theory are used to represent actors and links. A community is a set of actors and can be regarded as a cluster in a graph. A graph can be drawn as a network diagram. Communities are overlaid on the network diagram, and can be drawn as a closed area [7], an adjacent matrix [5], or others. A refined layout for the network can provide an intuitive and comprehensive image of networks.

We regard traditional work in exploring social networks as a network-centered approach. Most operations and information are attached to networks. These methods can give a general impression of entire networks and even an actor's most important properties. However, these methods pose two problems. The first is that images become cluttered since communities overlap, when actors are allowed to join more than one community. The second problem is the great visual burden imposed on the investigator in having to explore networks in detail because of the excessive amounts of information on the networks.

To solve these problems, we present an approach that uses a combination of three views corresponding to three aspects: nodes, links, and communities. Based on the approach, we developed a tool that integrates the network diagram, a tagcloud [8], and a list to support SNA. The network diagram shows links between actors and the whole network, the tagcloud displays actors with their attributes, and the list presents the communities that the actors belong to. We also provide a set of techniques of interaction to explore social networks.

2 Task Description

SNA aims at identifying important actors, crucial links, communities, and network characteristics [2]. We classified the tasks for SNA into three specific categories.

2.1 Task to Identify Actors, Links, and Communities

SNA should identify actors, links and communities with specific properties.

There are several properties of actors that an investigator is interested in. Of these, centrality is widely used to identify actors that occupy important positions. The representative criteria of centralities include degree centrality and betweenness centrality.

- Degree centrality evaluates the importance of an actor by the number of actors linked to it. An actor linked to most of the other actors is regarded as being the most important. An actor with high degree centrality can be interpreted as an outgoing or popular person in a friendship network.
- Betweenness centrality evaluates the importance of an actor by the number of times that the shortest path of any pair of other actors passes through the actor. The most mediated actor is considered to be the most important. The one with high betweenness centrality can be translated to be an intermediary of subgroups in a friendship network.

SNA should also identify crucial links such as those that bridge many pairs of actors, and communities of actors who have tight internal connections.

2.2 Task to Comprehend Details of Certain Actors, Links, and Communities

Details are also necessary for understanding the identified actor, links, and communities. SNA should provide investigators with actors that are directly linked to the interested actor and communities that the actor belongs to. SNA should also present links between actors in a way that is understandable. Basically, people want to know the two actors involved in a link, as well as actors linked to a certain actor. In addition, investigators want to know how actors in a community are connected, what the attributes of its members are, and who bridges multiple communities. SNA should comprehensively and clearly answer these substantive questions.

2.3 Task to Understand the Whole Network

SNA should provide an overview of the whole network. The datasets for SNA studies are not only about the attributes of an actor, but also systematic features of all the actors. These features reflect various social phenomena or affairs. SNA reveals these phenomena or affairs by providing an overview of the whole network.

3 Traditional Tools for SNA

There are a number of visualization systems used to support SNA. We divided these systems into two groups, i.e., statistical analysis-oriented and visual presentation-oriented methods.

Statistical analysis-oriented methods calculate statistical values that represent a variable feature of social networks. For example, UCINET [1] enables centrality to be measured, communities to be identified, and other functions for analysis. As UCINET has too many complex functions, users always need help from experts or need to study manuals. Pajek [3] is a tool embedded into UCINET. It enables users to understand social networks more easily by drawing the network as a diagram. However, it only depicts the statistical values in colors or the sizes of the nodes in a network diagram. When identifying actors and communities, users are very easily distracted by the complicated denotations in the network diagram.

Recent research has tended to take advantage of human vision, and provides visual presentations and various techniques of interaction to enhance exploration done by investigators. NodeTrix [5] integrates a network diagram and a matrix to enable the global structure of the network to be visualized and also to allow local communities to be analyzed in detail. However, it lacks the capacity for fine exploration of actors. This is because users have to take their time to distinguish an actor's centrality by assessing the brightness of nodes in a large network diagram. Perer et al. suggest a combinatorial view of a network diagram and a list [7]. However, it only provides dependent views of an actor list, while communities are drawn as closed areas and overlaid on the network. Their method only divides the network into disjointed subgroups. If these approaches were applied to displaying communities that all intersected without discrimination, a great number of overlaps would occur and result in visual clutter. As a result, it would be difficult to identify communities and content.

4 Our Approach: Views Corresponding to Three Aspects

We considered actors, links, and communities as three aspects from which investigators could start their explorations since these provide very distinct information. Each of these aspects has its own features and provides different information.

Our idea was to divide presentations of social networks into three appropriate views (see Fig. 1) that corresponded to the three aspects and synchronize these views to reveal other information that concerned aspects in real time. By doing so, information in networks can be dispersed in different views, so that investigators can rapidly and effectively identify all aspect by only focusing on one. In addition, other representations (e.g., lists) can fill in gaps where the network does not lay out networks sufficiently systematically such as in sorting. The following describes the three different views and the techniques of interaction linking them.

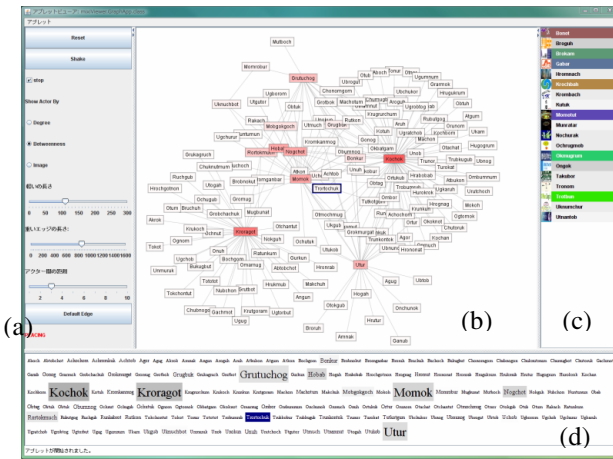


Fig. 1. Snapshot of MixVis. Actor clicked in tagcloud is highlighted by blue box in network. Actor’s communities are highlighted in different colors.

4.1 Tagcloud to Rapidly Identify Actors

The most important properties of an actor are its centrality, its absolute value, and how important it is in the whole network. We used a tagcloud to represent actors to identify and comprehend them with specified properties. The tagcloud is a visual presentation of a set of words. The attributes of text in a tagcloud such as size and color are used to represent features (such as frequency) of associated terms [8]. This is for “impression formation” and provides a pre-attentive image of the “key person”.

Actors’ names in our tagcloud were lined up on several lines. The layout of the actors was changed automatically according to the size of the tagcloud panel and the number of actors. The size and background of each actor was determined by its centralities. An actor’s size was proportional to its degree centrality and the brightness of its background was inversely proportional to its betweenness centrality.

The tagcloud also provided an accurate value for degree and betweenness centrality. When the cursor crossed over an actor, a label popped up, and the absolute values of degree and betweenness centrality were displayed on the label. The actors in a tagcloud could be sorted by their order of centrality and in alphabetical order. Fig. 1(d) shows a tagcloud with about 150 actors.

4.2 Network Diagram for Links and Overview of Networks

People are concerned about who is connected to whom in links, especially those actors connected to central actors. Other important information is in the overview of the network, which is the entire content for the links. Networks are drawn as network diagrams. We developed a force-oriented method to lay out the network diagram, by which we showed the links between actors, and also made it possible to overlay communities on the network diagram [4].

The method of drawing a compound graph was leveraged to draw the network. We defined a compound graph as a combination of an inclusion graph and an adjacency graph. It could not only represent links between actors, but also sets of actors (communities). We drew links as straight-line segments with our method, and drew communities as inclusion areas. Our method of drawing allowed communities to intersect. There is a drawing of a network in Fig. 1(b).

It is also possible to depict an actor's properties by using a node's attributes. This makes it convenient to check the link's features. For example, it is easy to know whether one link is connecting central actors or not, by checking the connected actors' colors. We designed three styles to draw nodes (actors). 1) Labels indicate actors' names and a label's background color represent an actor's degree centrality. A red background indicates that an actor's degree centrality is high (see Fig. 2). 2) Labels indicate actors' names and a label's background color represents an actor's betweenness centrality. A red background indicates actor's betweenness centrality is high (see Fig. 1). 3) An actor's image represents the actor.

The layout of the network can be refined to observe both the overview of the whole network and the details of one link. We provided a control panel (Fig. 1(a)), in which a slider was used to adjust several parameters to lay out the network diagram.

4.3 List to Flexibly Select Communities of Interest

A community indicates a set of actors sharing some common properties. Flexible interactions with communities are required to search and choose communities of interest. An investigator can comprehend the structure (pattern) of communities and find intersections through interactions. We used a list to select communities (see Fig. 1(c)).

The list contained all the communities actors belonged to. These communities were arranged in alphabetical order. An investigator could select single or multiple communities of interest.

The brightness of a community's background indicates its size, i.e., the number of members belonging to it. A dark community indicates that it has many members. When highlighting communities, e.g., when multiple communities are selected, we use another set of colors to distinguish them. We used a hue, saturation, and value (HSV) color model to assign colors to the communities to avoid confusion. More

importantly, we produced different bright colors, rather than gray, by controlling the hue, saturation, and value.

4.4 Linking Three Views

The network diagram, tagcloud, and list were not isolated views, but an integrated, synchronized system. We introduced the concept of “linking and brushing” to design our system. Linking and brushing is a technique to combine different representations to reveal different aspects of data. Interaction with one will cause corresponding changes in the other representations.

We leveraged the technique to produce insights into each aspect that was concurrently visible, so that the investigator could rapidly acquire details about the aspect of concern. We mainly harmonized the three views by highlighting related information in the two other views when selecting an object in the one view. Moreover, we used the same color to highlight a community and its members. By doing so, the investigator could visually distinguish members and their affiliations.

5 MixVis: A Tool for Exploring Social Networks

Our system is constructed of three main views including a network diagram, a list, and a tagcloud. The middle panel (Fig. 1(b)) is the network diagram, the left of which has the control panel (Fig. 1(a)) that contains widgets to adjust the diagram. The community list (Fig 1(c)) is at the right of the network diagram. A tagcloud (Fig 1(d)) is at the bottom of the whole user interface.

The size of the view can be adjusted by users themselves according to their individual needs. They can also hide or maximize one view. In some cases, an appropriate size can enhance identification. For example, an extended network diagram can clearly show large-scale nodes and their relationships.

Our tool provides techniques for interactively exploring multiple aspects of insights into social networks. These insights contain an overview of the whole network, insights into communities or parts of the network, and details on some actors or links. They are consistent with people’s interests, i.e., properties of a single actor, links, communities, connectivities between or intersections of communities, and the balance or density of the whole network. In the following, we will describe techniques of interaction with MixVis.

5.1 Interactions from Tagcloud

We assumed that an investigator would be focusing on actors and want to find those who were central. Tagcloud shows actors in an impressive way in that: 1) actors with high degree centrality are outstanding among others since they are larger and 2) actors with high betweenness centrality are easy to find because their backgrounds are darker. Consequently, an investigator can find them at a quick glance in a tagcloud.

Moreover, they can accurately identify central actors by sorting them by degree centrality, betweenness centrality, or in alphabetical order. Popup labels that post the

value of centrality can facilitate this process. After this, the investigator can know the position of the actor in the network and its communities by clicking a tag. The same actor will be highlighted in the network and its communities will also be highlighted.

Fig. 1 illustrates how the other two views change after the tag is clicked. One tag is identified and selected. Due to the three views being linked, the actor is also highlighted by a box in the network diagram and the communities that the actor belongs to are highlighted by different colors on a list.

5.2 Interactions from Network Diagram

When browsing the network diagram, an investigator may notice an actor and want to find out whom it is related to, which community it belongs to, and what its centrality is. The role of the actor is investigated by clicking it in the network. Actors directly connected to the one that was clicked are highlighted by a red box, while the other actors are blurred. Communities that the clicked actor belongs to are highlighted on a list by the color of the background. The actor is also highlighted in a tagcloud, where we can see its other properties. The investigator can adjust the layout of the network diagram by using the widgets in the control panel.

Fig. 2 is snapshot of MixVis after one actor in the network has been clicked. By taking a glance at the three views, we can know its links, its properties such as its degree centrality, and its communities.

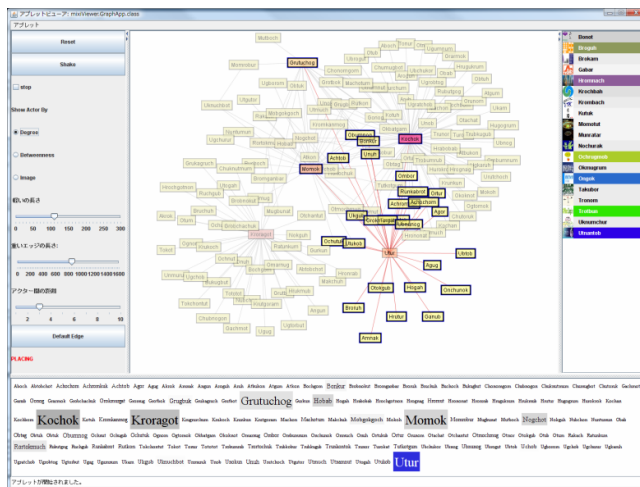


Fig. 2. Actor clicked in network diagram is turned blue in tagcloud. Its communities are highlighted in different colors.

5.3 Interactions from List

If an investigator focuses attention on communities, the list can facilitate his/her work by allowing any of one or more communities to be selected. In the meantime, the selected communities are drawn as inclusion areas in the network. As actors shared by multiple

communities are placed at the intersection of closed areas, the investigator can find them easily. To see the centralities of the members of every community, tags belonging to these communities are highlighted in the same color as the respective community.

As seen in Fig. 3, two communities on the community list have been selected; the actors are contained in an inclusion area to indicate their membership. The correspondence is very clear from the color of the community, the inclusion area, and the actor in the tagcloud. Therefore, that investigator can easily comprehend the details of the communities.

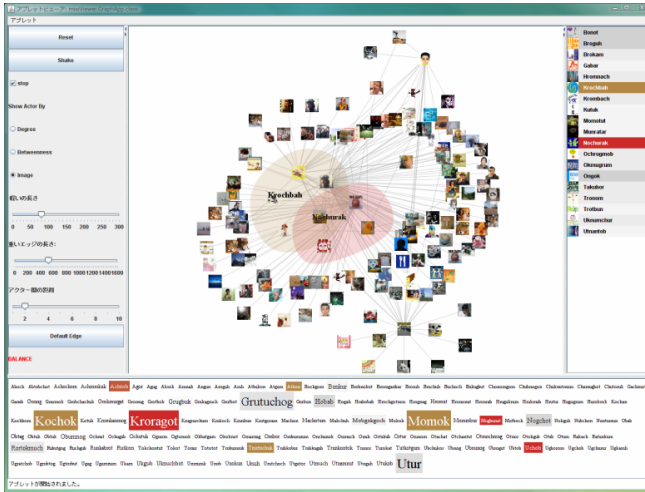


Fig. 3. Two communities are selected. Colors of members in tagcloud area in network diagram are consistent with communities on list.

6 Evaluation

6.1 User Study

To test and verify the efficiency of our approach, we administered a user study in which MixVis was compared to two other tools: ToolA and ToolB. ToolA only had the network diagram in MixVis, and the other information related to the community and actor was overlaid on the network diagram. ToolA was similar to the traditional tools that presented all the information in a network diagram. ToolB included a network diagram and a tagcloud. The list in ToolA and ToolB could be displayed or hidden all together. ToolB was similar to work that also used combinations of different representations.

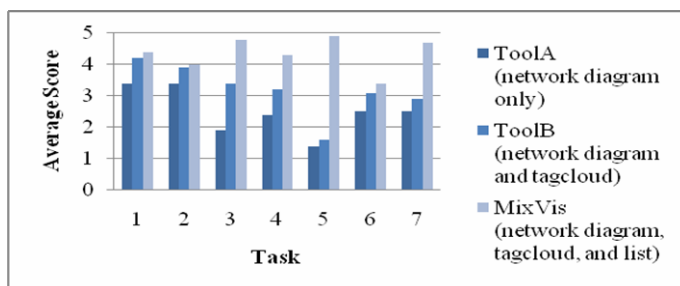
We designed seven tasks and assumed them to be important tasks of SNA. There is a brief description of the tasks in Table 1. The tasks covered the categories given in Section 2.

Table 1. Tasks to be solved by subjects

Task No.	Content
1	To identify actor with highest degree centrality
2	To identify actor with highest betweenness centrality
3	To explore details on an assigned actor
4	To explore details on a specified link
5	To explore details on two specified communities
6	To find groups of actors by browsing the whole network
7	To recognize a trend related to the community and actor by using three views

We asked ten students to participate in the user study. They either knew little or a great deal about visualization. Prior to starting this study, we gave them a short introduction to SNA and the functions of three tools. Then, they were asked to solve the seven tasks by using ToolA, ToolB, and MixVis (To eliminate the influence of names, we assigned the new name of “ToolC” to MixVis when conducting the study). After each task, subjects awarded these three tools scores according to how usable they were. Five points represented the maximum score.

The average scores for the three tools for each task are shown in the bar graph in Fig. 4. The figure indicates that MixVis was awarded the highest score for handling the seven tasks.

**Fig. 4.** Average scores for ToolA, ToolB, and MixVis

6.2 Discussion

According to Fig. 4, we assessed that MixVis was better than the other two tools in usability for most of the tasks.

The differences in the average scores strongly support our perspective that the representation of multiple aspects can facilitate an investigator’s work. Significant differences, especially in completing Tasks 5 and 7, appeared and were verified with a t-test. We inferred that these differences were caused by the fact that the two tasks were related to the community and only MixVis could provide a list and a flexible function to enable selection.

ToolA, ToolB, and MixVis were evaluated similarly in Tasks 1 and 2. The t-test confirmed that there are no significant differences between ToolA and MixVis in accomplishing Task 2. There were also no significant differences between ToolB and

MixVis in accomplishing Task 1. The reason for the lack of difference is that in Tasks 1 and 2, we asked subjects to identify the actor who had the highest degree centrality and highest betweenness centrality that could be done without using a tagcloud and a list as in the traditional method.

7 Conclusion

We proposed an approach that used a combination of three views corresponding to nodes, links, and communities. Based on the approach, we developed a tool called MixVis to support SNA. We adopted “linking and brushing” to enable real-time interaction between the three different views, and interaction between users and the proposed tool. Our tool allowed drawings of an actor’s affiliations, which is not supported by most traditional work. We administered user studies to evaluate how usable our tool was. Another two tools, which were equipped with only one or two views, were compared to the new tool. The experimental results revealed that MixVis was much easier to use.

References

1. Borgatti, S., Everett, M., Freeman, L.: UCINET V user’s guide. Analytic Technologies (1999)
2. Brandes, U., Wagner, D.: Visone - Analysis and Visualization of Social Networks. In: Jünger, M., Mutzel, P. (eds.) Graph Drawing Software, pp. 321–340. Springer, Heidelberg (2004)
3. de Nooy, W., Mrvar, A., Batagelj, V.: Exploratory Social Network Analysis with Pajek. In: Structural Analysis in the Social Sciences. Cambridge University Press, New York (2005)
4. Gao, J., Misue, K., Tanaka, J.: Drawings of compound graph using free-form curves. In: Proceedings of the 70th National Convention of IPSJ, vol. 1, pp. 407–408 (2008)
5. Henry, N., Fekete, J.-D., McGuffin, M.J.: NodeTrix: A Hybrid Visualization of Social Networks. *IEEE Transaction on Visualization and Computer Graphics* 13(6), 1302–1309 (2007)
6. Newman, M.E.J.: Fast algorithm for detecting community structure in networks. *Physical Review E (Statistical, Nonlinear, and Soft Matter Physics)* 69(6) (2004)
7. Perer, A., Shneiderman, B.: Balancing Systematic and Flexible Exploration of Social Networks. *IEEE Transaction on Visualization and Computer Graphics* 12(5), 693–700 (2006)
8. Rivadeneira, A.W., Gruen, D.M., Muller, M.J., Millen, D.R.: Getting our head in the clouds: Toward evaluation studies of tagclouds. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI 2007, pp. 995–998. ACM, New York (2007)