Layout Effects on Sociogram Perception

Weidong Huang[1,2], Seok-Hee Hong[1,2], and Peter Eades[1,2]

[1] IMAGEN Program, National ICT Australia Ltd.
[2] School of Information Technologies, University of Sydney, Australia
{weidong.huang, seokhee.hong, peter.eades}@nicta.com.au

Abstract. This paper describes a within-subjects experiment in which we compare the relative effectiveness of five sociogram drawing conventions in communicating underlying network substance, based on user task performance and usability preference, in order to examine effects of different spatial layout formats on human sociogram perception. We also explore the impact of edge crossings, a widely accepted readability aesthetic. Subjective data were gathered based on the methodology of Purchase et al. [14]. Objective data were collected through an online system.

We found that both edge crossings and conventions pose significant affects on user preference and task performance of finding groups, but either has little impact on the perception of actor status. On the other hand, the node positioning and angular resolution might be more important in perceiving actor status. In visualizing social networks, it is important to note that the techniques that are highly preferred by users do not necessarily lead to best task performance.

1 Introduction

Social networks can be modeled as graphs, and visualized as node-edge diagrams where nodes represent *actors*, and edges represent *relationships* between them. With advances in display media, the use of node-edge diagrams or *sociograms* (see Figure 1 for an example) has been increasingly important and popular in social network analysis. Sociograms serve as simple visual illustrations in helping people to explore and understand network structure, and to communicate specific information about network characteristics to others.

One of the major concerns in network visualization is *effectiveness*. There are two issues involved here: one is readability. Readability can be affected by not only intrinsic network characteristics [6], but also layout. In particular, *edge crossings* has long been widely accepted a major aesthetic [13, 4]. Purchase [16], in her pioneering work of a user study which compared the relative effects of five aesthetic criteria (bends, crosses, angels, orthogonality and symmetry) on abstract graphs, also concluded that edge crossings has greatest impact on human graph understanding. Subsequently the aesthetic of edge crossings was validated on UML diagrams [14]. The remark of Purchase et al. [14] that there is no guarantee that results of domain-independent experiments could automatically apply to domain-specific diagrams motivates this paper. We are not aware of any previous user studies in examining edge crossings impact on the perception of social networks when domain-specific tasks are performed.

P. Healy and N.S. Nikolov (Eds.): GD 2005, LNCS 3843, pp. 262–273, 2005.

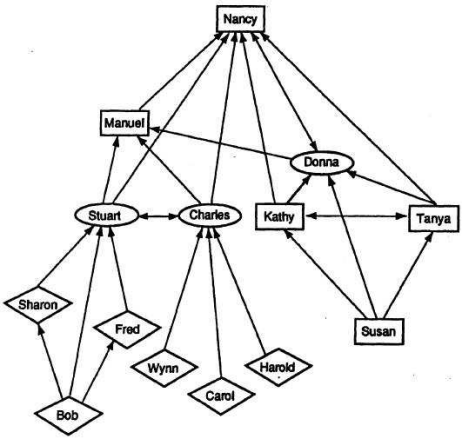

Fig. 1. Advice network formed by an auditing team. Courtesy of Krackhardt [9]. Ellipses represent managers; diamonds represent staff auditors and boxes represent secretaries. A line from Donna to Nancy indicates that Donna seeks advice from Nancy.

The other is communication. Good readability does not necessarily lead to effective underlying semantics communication. When a social network is mapped to a sociogram, what matters is relationship patterns, not the physical positioning of nodes [17]. However, previous studies [10, 11] revealed that the spatial layout of nodes in a sociogram does affect viewers in perceiving social network characteristics. Therefore there is practical need for investigating the actual communication effectiveness of a particular visualization method. Surprisingly, although considerable amount of fancy techniques have been proposed in literature, very little empirical evidence is available to support their effectiveness in communicating network structure to humans [12].

To address the above questions, we conducted a user study. In this study, we compared the communication effectiveness of five sociogram drawing conventions, and investigated the impact of edge crossings under each convention, based on user preference and task performance. Subjective data were gathered based on the methodology of Purchase et al. [14]. Objective data were collected through an online system.

1.1 Sociogram Constructing Conventions

Many visualization techniques are aimed to highlight one or two aspects of the network structure, and confirm to some aesthetics to improve the readability. Of our particular interest are the following five conventions.

1. Circular layout: all nodes are placed on a circle [17].
2. Hierarchical layout: nodes are arranged by mapping actors' status scores to the nodes' vertical coordinates [2].
3. Radial Layout: all nodes are laid on circumference of circles in a way that their distances from the center exactly reflect their centrality levels [3].

4. Group layout: nodes are separated into different groups with nodes in the same group close to one another. See [5] for a review.
5. Free layout: nodes are arranged without any particular purpose.

For more details about drawing conventions and background, see [7].

2 Experiment

2.1 Subjects

Twenty-three subjects were recruited from a student population in computer science on a completely voluntary basis. All the subjects were postgraduates and computer literate. All had node-edge diagram experience such as UML or ER, associated with their study units; six of them were graph drawing research students. All had neither academic nor working experience related to social networks. They were reimbursed $20 each for their time upon the completion of their tasks.

2.2 Design

Networks. For this experiment, two networks are used. One is Krackhardt's *advice network* [9]. The network is modeled as a directed graph with 14 nodes and 23 edges as shown in Figure 1. The network has three groups in a sense that we discuss later in this section.

The Katz status scores [8] of the actors are shown in Table 1. In this study, Katz status score was used as the index of importance. We expected subjects to perceive the importance in accordance with, and we measured their perceptions against the actors' Katz status scores.

The second one is a fictionalized network which was produced from the first one by eliminating all the directions. This gives an undirected graph with 14 nodes and 23 edges, which we call a *collaboration network*. A line between A and B means that A and B collaborate with each other.

Sociograms. We used a total of 12 drawings: 1) for each of the five conventions, two drawings of the advice network - one with minimum crossings and one with many crossings (see Figure 2); 2) for the collaboration network, a free convention drawing with minimum crossings and a free convention drawing with many crossings.

Table 1. Actors' Katz status scores

Nancy	1.00	Fred	0.02
Donna	0.66	Sharon	0.02
Manuel	0.57	Harold	0.00
Stuart	0.19	Wynn	0.00
Charles	0.17	Susan	0.00
Kathy	0.08	Bob	0.00
Tanya	0.08	Carol	0.00

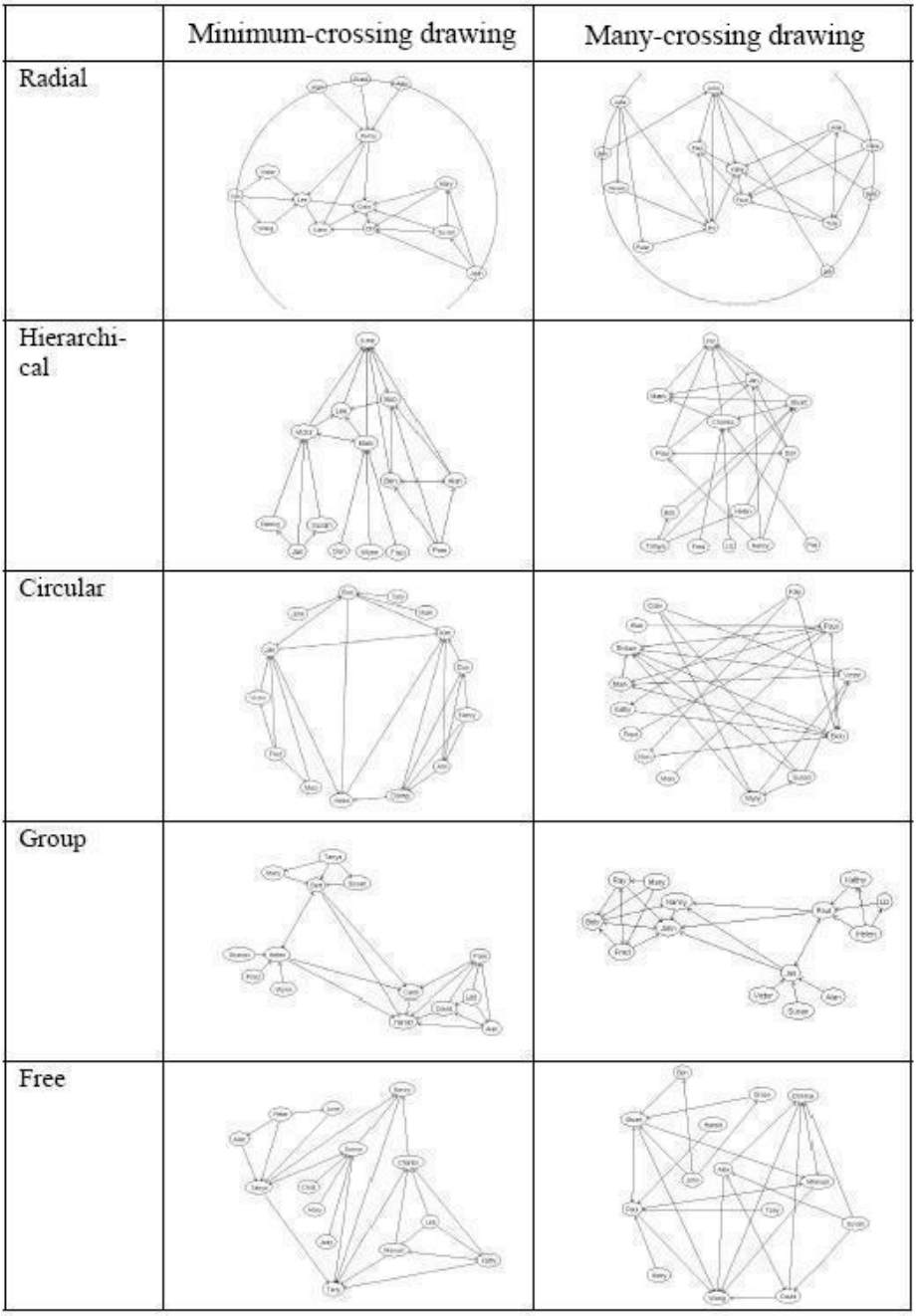

Fig. 2. Sociograms for the advice network used in the study

All nodes were labeled with different names; in every drawing, each node was mapped to a new name. By providing a context and background for each network, and names for actors, subjects were expected to perform tasks from the real world social network perspective [10]. However, subjects were not made aware that the drawings had the same graph structure.

Tasks. For this experiment, we considered two common social network measures which are frequently highlighted in sociograms: one is *importance* or status of actors; the other is the *presence of social groups*, in which the connections among actors are relatively dense. The whole session included 3 main tasks:

1. Online tasks
   (a) find 3 most important actors and rate them according to their importance levels; and
   (b) determine how many groups are in the network, and separate the 4 highlighted actors according to their group membership, given the condition that one actor should not belong to more than one group, and one group should not include only one actor. In formal tests, the same 4 nodes (actors) across all drawings were highlighted in red rectangles.
2. Subjective rating tasks
   (a) Usability acceptance rating: with one page showing all 6 many-crossing drawings, and the other page showing all 6 minimum-crossing drawings, subjects were required to rate their usability based on a scale from -3 (completely unacceptable) to +3 (completely acceptable) for importance tasks and group tasks, respectively.
   (b) Crossing preference rating: each many-crossing (A) and minimum-crossing (B) pair was shown one by one, 6 pairs in total. Subjects needed to indicate their preferences for importance tasks and group tasks, respectively, based on a scale from -2 (strongly A) to +2 (strongly B), where, for example, "Strongly A" means A is strongly preferred over B.
   (c) Overall usability ranking: with all 10 advice network drawings being shown in one page, subjects needed to choose 3 drawings that they least preferred and 3 drawings that they most preferred for their overall usability, using a scale from -3 to -1 and from 1 to 3, respectively.
3. Questionnaires: there were 2 questionnaires with each having a different focus, and to be presented to subjects before and after they were debriefed about edge crossings and drawing conventions, respectively. The first questionnaire asked subjects information about their study background, experience with node-edge diagrams and social networks, how they interpret sociograms, and any network structure and sociogram features that they think may influence their graph perceptions. The second questionnaire asked about their thoughts about conventions and edge crossings.

For the above rating tasks, subjects were also required to write down a short explanation for each answer.

Online System. Sociograms were displayed by a custom-built online experimental system. The system was designed so that:

- A question is shown first, a button on the screen is pressed, then the corresponding drawing is shown; after writing down the answer on the answer sheet provided, the button is pressed and the next question is shown, and so on.
- Each subject's response time for each drawing is logged. This starts once a drawing is completely displayed and ends once the button is pressed.

The study employed a within-subjects design. For online tasks, 10 of the 12 drawings were randomly chosen and shown to comply with the time schedule and to reduce fatigue. The order of group and importance questions for each drawing was also random. Subjects were told they could have breaks during the question viewing periods if they wished. There was no time limit on task completion, although they were recommended to answer each question in one minute. During the preparation time, subjects were instructed to answer each question in the context of the underlying network and as quickly as possible without compromising accuracy.

2.3 Procedure

A pilot study had been conducted with another four subjects who did not have any social network background to check our methodology. They showed that they quickly understood the questions and felt comfortable with the experiment. They related the visual network representations with their daily social experiences when performing tasks.

The formal tests took place in a computer laboratory, in which all PCs had the same specifications. Before starting the experiment, subjects were asked to read the information sheet, sign the consent form, read through and understand the tutorial material, ask questions and practice with the online system. The drawings used for practice were quite different from the ones used in formal tests, since the practice was only for familiarization with the procedure and system, not for them to get experienced with sociogram reading.

Once ready to start, subjects indicated to the experimenter, and started running the online system performing tasks formally. After the online reading tasks followed by a short break which was to refresh subjects' memory, they proceeded with the rating tasks, and the first questionnaire. Then, after being given a debriefing document explaining the nature of the study, edge crossings and drawing conventions, subjects were asked to do the rating tasks (a) and (b) again, and finally finished with the second questionnaire. Subjects were also encouraged to verbalize any thoughts and feelings about the experiment. The whole session took about 60 minutes.

3 Results

The data of three subjects were discarded due to the failure of following instructions. Since the collaboration network sociograms did not create distinct difference with their counterparts, these data have been omitted in our analysis.

For simplicity, we use C, R, G, H, and F to represent **c**ircular, **r**adial, **g**roup, **h**ierarchical and **f**ree drawing conventions, respectively, and use P for minimum-crossing drawing, and C for many-crossing drawing. Therefore CP denotes circular minimum-crossing drawing; CC denotes circular many-crossing drawing, and so on.

3.1 User Preference Data

Usability Acceptance. Subjects' usability rating scale data are illustrated in Figures 3-4 and analyzed using ANOVA (with Fisher's PLSD, for pairwise comparisons).

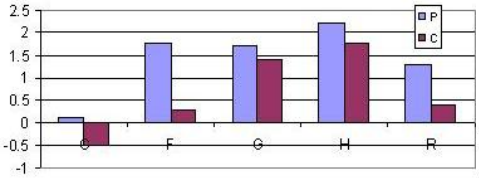

Fig. 3. Mean importance usability scores

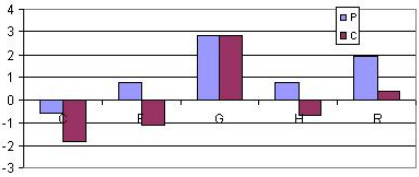

Fig. 4. Mean group usability scores

Table 2. Pre- and post-debriefing scores (only data with significant changes are shown)

Importance	HP	HC	RP	RC	Group	RP	CC	FC	HC
Pre	2.21	1.75	1.29	0.42	Pre	1.96	-1.83	-1.13	-0.67
Post	2.75	2.29	1.87	1.63	Post	1.50	-1.43	-0.54	-0.08
P Value	0.04	0.05	0.00	0.00	P Value	0.02	0.03	0.02	0.02

For importance tasks, the minimum-crossing drawing was generally rated higher than the many-crossing one for each convention. There was a significant crossings effect between all minimum-crossing and all many-crossing drawings (p=0.00). Also there was a significant convention effect among all drawings (p=0.00). The pairwise comparisons revealed that the user acceptance for CP was significantly different for GP, HP and FP, respectively; the user acceptance for RP was significantly different for CP and HP, respectively. Furthermore, *paired t tests* showed that after debriefing, the drawings of both hierarchical and radial conventions were rated significantly higher (see Table 2). In particular, the mean scores of the hierarchical convention pair were higher than all others, showing that the positioning of nodes was perceived more important than edge crossings for importance tasks.

For group tasks, GC and GP were rated much higher than others; in fact the others were perceived as having little usefulness. Analysis found a significant crossings affect for each pair of drawings for all conventions (p=0.00) except group convention. Also, conventions produced a significant difference among all drawings (p=0.00). Pairwise comparisons showed that all pairs were significantly different except between HP and FP. In subjects' post-debriefing ratings, there was a significant change for RP, CC, FC and HC (see Table 2).

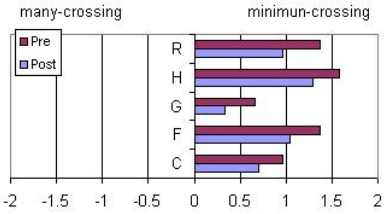

Fig. 5. Mean importance preference scores

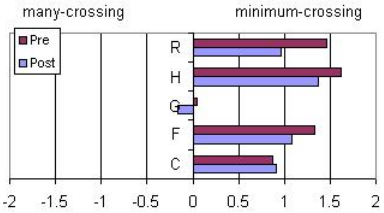

Fig. 6. Mean group preference scores

Crossing Preference. As can be seen from Figures 5-6, generally, subjects preferred the minimum-crossing drawing more for each convention. The *1-sample t tests* against the hypothesized mean (=0) revealed that for all conventions except group convention, subjects' preference for the minimum-crossing drawings over the corresponding many-crossing drawings was statistically significant ($p<0.01$).

The *paired t tests* revealed that there were no significant changes between the pre- and post-debriefing ratings, although the post-debriefing preferences were generally weaker than pre-debriefing preferences for both importance and group tasks.

Overall Usability Ranking. Subjects' ranking values for each drawing were summed as a weighted value, and weighted values for all drawings are shown in Figure 7. It can be seen that generally, the many-crossing drawings were less preferred except GC, which was ranked the highest for its overall usability, followed by GP, then HP. Both CC and FC had the lowest weighted value, indicating that they were considered having little overall practical utility.

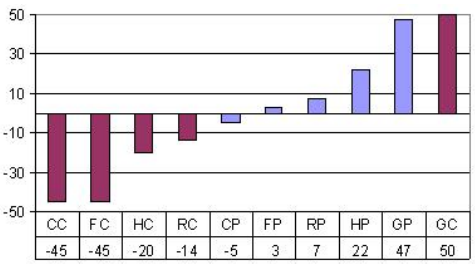

Fig. 7. Overall usability values

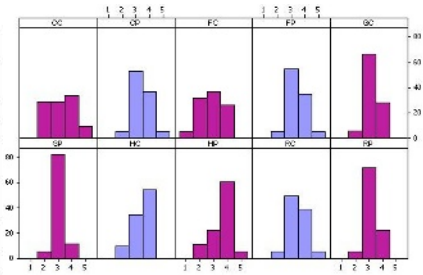

Fig. 8. Distributions of reported group number

3.2 User Performance Data

Response Time. Subjects' response time data are illustrated in Figures 9-10 and analyzed using the non-parametric method of Kruskal-Wallis.

For importance tasks, subjects spent shorter time with the minimum-crossing drawing than the many-crossing drawing for each convention in general. Among all minimum-crossing drawings, the shortest time was spent with GP, followed

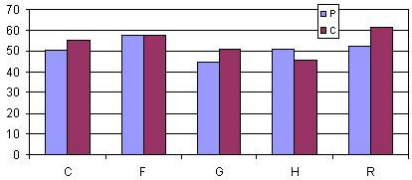

Fig. 9. Median time (sec.) for importance

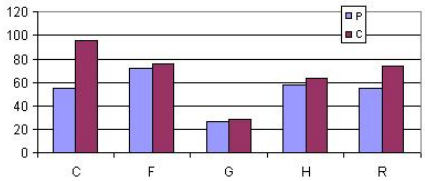

Fig. 10. Median time (sec.) for group

by CP, HP, RP, and finally FP. However, statistical tests revealed that these differences of response time were not statistically significant in terms of the effect of either edge crossings or drawing conventions.

For group tasks, following a similar pattern, again shorter time was spent with the minimum-crossing drawing than the many-crossing one for each convention. Among all minimum-crossing drawings, the shortest time was spent with GP, followed by RP, CP, HP, and finally FP. Analysis showed that there was a significant crossings effect between the minimum-crossing and many-crossing pair of circular convention (p=0.012), and between all minimum-crossing and all many-crossing drawings (p=0.021), and a significant convention difference among all drawings (p=0.000). Pairwise comparisons showed that subjects spent significantly shorter time with GP than all others at the level of 0.01.

Reported Group Number and Member Group Assignment. Figure 8 illustrates the distribution of the reported group number for each drawing. As can be seen, GP had largest proportion of subjects (82.4%) responded "correctly"(3 as expected). An analysis of variance of the reported group number for all drawings showed there was a significant difference at the level of 0.066.

Also, at dyad level, the member group assignment task was to investigate edge crossings and convention impact on the perception of actors' co-memberships. As can be seen from Figure 11, a relatively larger proportion of subjects performed this task correctly on the minimum-crossing drawing than on the many-crossing drawing for each convention except free convention. Among all minimum-crossing drawings, GP yielded the highest correctness rate (76.5%).

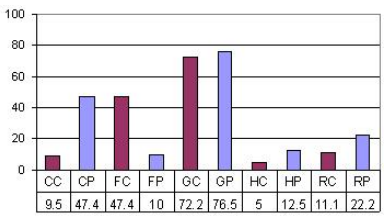

Fig. 11. Group assignment correctness(%)

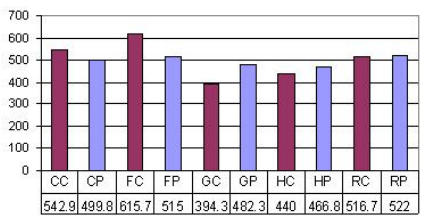

Fig. 12. Weighted values for identifying important actors

Identifying Most Important Actors. Figure 12 shows the weighted values for all drawings. The weighted value is to measure a drawing's overall effectiveness of conveying information about importance, and calculated in the following way: First we gave an index of 5 to the most important actor, 2 to the second and 1 to the third; then the productions of indices and corresponding correctness percentages were summed as a weighted value for each drawing. It can be clearly seen that FC had the highest weighted value, followed by CC, RP, and then RC; GC had the lowest.

4 Qualitative Data and Discussions

The analysis of subjects' responses to questionnaires and interviews revealed that subjects had a strong preference of placing nodes on the top or in the center to highlight importance, and clustering nodes in the same group and separating groups to highlight groups. They had tendency to believe that nodes in the center or on the top are more important, and nodes in close proximity belong to the same group.

There was strong evidence that edge crossings contribute to the significant difference in user preference, usability acceptance, and group task performance. Edge crossings not only affect the ease of reading, but also affect the understanding of network structures.

With respect to drawing convention, for importance tasks, hierarchical convention was strongly preferred, while for group tasks, group convention was strongly preferred. Users achieved the highest response accuracy with group convention for group tasks. However, the highest response accuracy did not come with hierarchical convention for importance tasks. For overall usability, group convention was the one for which the usability was rated high and user performance was well as well.

Quite surprisingly, subjects were overwhelmingly in favor of hierarchical convention for importance tasks; they spent relatively short time with HP, but obtained the lowest correctness rate among all minimum-crossing drawings. On the other hand, FC obtained the highest correctness rate, but relatively long time was spent with it. We realized that some subjects had complained in questionnaires and verbally that in some drawings, edges were incident to nodes too closely to clearly identify arrow directions. Visual inspection revealed that indeed, free convention drawings had very good angular resolution, while hierarchical convention made angular resolution relatively low, where edges had to be crowded in one side of nodes. In addition, subjects spent longer time with FC, which might actually allow them to have better chance to understand network structure better.

In summary, no obvious evidence was found that either edge crossings or conventions pose significant impact on user importance task performance. Users generally performed better when they took longer time. We conjecture that only those tasks which are closely related to edges and involve edge tracing can be significantly affected, such as finding groupings. On the other hand, for

communicating information about actor status, the angular resolution and node positioning in a sociogram might be more important, compared to drawing conventions and reducing the number of edge crossings.

For detailed discussion of user responses, the recommendations for sociogram design, and some further hypotheses derived about human graph perception, see [7].

5 Conclusions

This study, together with previous research [10, 11], has demonstrated that how sensitive the human sociogram perception is to spatial layout, and how important it is to have visualization techniques evaluated for their actual effectiveness in communication from human understanding point of view. It should be noted that visualization techniques, which are highly preferred by users, do not necessarily always produce best task performance, as demonstrated in this study.

The findings from this study should be interpreted within the limitations of the given experimental settings. In this study we had only investigated the relative effectiveness of five "explanatory visualization" [2] conventions and edge crossings impact under each convention, in communicating actor status and subgroup information to novice audience. Their usability in assisting professionals to explore and understand social networks remains untouched and is beyond the scope of this study. For a comprehensive overview in this field, see [1].

Additional studies are needed to empirically identify and investigate the impact of other possible variables of human sociogram perception in a more controllable manner.

Acknowledgements. The authors are grateful to Mr Kelvin Cheng, and Mr Le Song for their helpful comments on questionnaire design, and to students who willingly took part in the experiment. Ethical clearance for this study was granted by the University of Sydney, December 2004.

References

1. Card, S., Mackinlay, J., and Shneiderman, B. (1999): Readings in information visualization: using vision to think. Morgan Kaufmann Publishers, San Francisco, CA.
2. Brandes, U., Raab, J. and Wagner. D. (2001): Exploratory Network Visualization: Simultaneous Display of Actor Status and Connections. Journal of Social Structure 2(4).
3. Brandes, U., Kenis, P. and Wagner, D. (2003): Communicating Centrality in Policy Network Drawings. IEEE Trans. on Visualization and Computer Graphics 9(2): 241-253.
4. Di Battista, G., Eades, P., Tamassia, R. and Tollis, I. (1998): Graph drawing: algorithms for the visualisation of graphs, Prentice Hall.
5. Freeman, L. (1999): Visualizing Social Groups. American Statistical Association, Proceedings of the Section on Statistical Graphics, 47-54.

6. Ghoniem, M., Fekete, J. and Castagliola, P. (2004): A Comparison of the Readability of Graphs Using Node-Link and Matrix-Based Representations. Proceedings of the 10th IEEE InfoVis'04, Austin, TX. IEEE Press. pp. 17-24.
7. Huang, W., Hong, S., and Eades, P. (2005): Layout Effects: Comparison of Sociogram Drawing Conventions. http://www.it.usyd.edu.au/∼whua5569/ex/.
8. Katz, L. (1953): A new status index derived from sociometric analysis. Psychometrika, 18: 39-43.
9. Krackhardt, D. (1996): Social Networks and Liability of Newness for Managers. In C. L. Cooper and D. M. Rousseau (eds.), John Wiley Sons, Ltd. New York, NY. Trends in Organizational Behavior, Volume 3, pp. 159-173.
10. McGrath, C., Blythe, J. and Krackhardt, D. (1997): The effect of spatial arrangement on judgments and errors in interpreting graphs. Social Networks, 19(3):223-242.
11. McGrath, C., Blythe, J. and Krackhardt, D. (1996): Seeing Groups in Graph Layout. Connections, 19(2): 22-29.
12. McGrath, C., Krackhardt, D. and Blythe, J. (2003): Visualizing Complexity in Networks: Seeing Both the Forest and the Trees. Connections, 25(1): 37-47.
13. Moreno, J. L. (1953): Who shall survive: Foundations of Sociometry, Group Psychother-apy, and Sociodrama. Beacon House Inc.
14. Purchase, H., Allder, J. Carrington, D. (2002): Graph Layout Aesthetics in UML Diagrams: User Preferences. J. Graph Algorithms Appl. 6(3): 255-279.
15. Purchase, H., Cohen R., James M. (1995): Validating graph drawing aesthetics. Proceedings of the Graph Drawing Symposium, Passau, Germany (GD '95), 435-446, Springer-Verlag.
16. Purchase, H. (1997): Which aesthetic has the greatest effect on human understanding? Proceedings of the 5th International Symposium on Graph Drawing (GD '97), 248-261. Springer-Verlag.
17. Scott, J. (2000): Social Network Analysis: A Handbook. Sage Publications, 2nd edition.