# How People Read Graphs

**Weidong Huang, Peter Eades**

The University of Sydney
National ICT Australia

whua5569@usyd.edu.au, peter.eades@nicta.com.au

## Abstract

The graph layout problem has long been a major concern for effectiveness of conveying information. To propose user-centred aesthetic criteria for a "good" layout, it is important to have knowledge on how people read graphs; how a particular graph layout characteristic can affect people's reading performance. On the other hand, despite the increasingly wide use of graphs in everyday life, yet we know surprisingly little about how people actually read graphs. The present eye tracking study in this paper is an attempt to perform an initial investigation into this issue and provide data that can help build the basic understanding of how people read graphs.

*Keywords*: graph reading, graph drawing, edge crossing, eye movement, eye tracking.

## 1 Introduction

The use of graphs has been increasingly important in our daily lives. Many real world problem domains can be modeled by graphs and visualized in the form of a node-edge diagram where nodes represent objects while edges represent the relationships between the objects (Colin Ware et al 2004). Since graphs are drawn to illuminate application data, and different graph layouts can differ dramatically in effectiveness of conveying the information embodied in the graph to viewers, the layout problem has long been a major concern in graph drawing research.

Some aesthetic criteria have been proposed for this purpose for two dimensional graphs. Examples of these criteria include symmetry (e.g. Peter Eades 1984, R. Lipton et al 1985), minimum edge crossings (e.g. D. Ferrari et al 1969), and minimum bends (e.g. H. Trickey 1988, R. Tamassia 1987). Unfortunately, these criteria were not originally based on experimental data. While these criteria have yielded some valuable insights in terms of people understanding graphs (e.g. Colin Ware et al 2002, Helen Purchase et al 1995), there might be other better criteria, or graph features which are not yet found, but can play an important role in graph understanding. To define user-oriented graph layout criteria, it is important for us to have knowledge about how people read graphs first.

Copyright © 2005, Australian Computer Society, Inc. This paper appeared at *Asia Pacific Symposium on Information Visualisation (APVIS 2005)*, Sydney, Australia. Conferences in Research and Practice in Information Technology, Vol. 45. Seok-Hee Hong, Ed. Reproduction for academic, not-for profit purposes permitted provided this text is included.

However, despite the popularity of graphic communication being continuously on the increase, relatively little is known about how people extract information from graphs. In general, the attention paid to graphs by researchers has been disproportionately small relative to the extent of their use (Aidan Feeney et al 2000). Although, recently much work has been done on people viewing statistical graphs (such as histograms), maps, and images (e.g. Aidan Feeney et al 2000, Carpenter, P. A. et al 1998, Jason S. et al 2003, Raj M. et al 2003), yet we know little about how people read graphs represented by node-edge diagrams.

On the other hand, the remarkable work of Helen Purchase (1997) and Colin Ware et al (2002) has revealed that some drawing properties have important impact on graph understanding. For example, the number of edges that cross the shortest path, and the continuity of multi-edge paths (Colin Ware et al 2002). However, their observations were mainly based on the response time and correctness rate, and cannot therefore explain where the time is spent and how the correctness rate is affected by these properties.

During picture viewing, eye movements are not random, rather guided toward informative and interesting regions and targets (Buswell, G. T. 1935, Yarbus, A. L. 1967). An eye tracking experiment was carried out based on the above motivations. Our primary interest in this study is in how people execute visual queries when reading graphs. To be more specific, by examining the eye movement data recorded by an eye tracking system, we hope to discover some relationships between eye movement patterns and properties of graph drawings.

One point that is worthy to be mentioned here is that, in practice, graph reading can be open-ended or aim-oriented. However, from the graph drawing point of view, the assumption we made for this study is that graph reading is an aim-oriented task, based on the observation that at most cases, people view graphs in order to obtain specific information from a particular drawing.

## 2 Experiment

### 2.1 Participants

Thirteen participants were recruited from student population on completely voluntary basis. Those with glasses and contact lenses were not further considered, due to limitations of the eye tracker. All the participants are postgraduates who have normal naked vision, and completely new to eye tracking. Three of them had intensive graph reading experience, two had no experience at all, and others had limited experience only related to database or/and information visualization courses. They were reimbursed $20 each for their time upon the completion of their tasks.

## 2.2 Apparatus

The study was conducted in the Pervasive Lab of the Information Technologies School at The University of Sydney. The testing room contains one operator PC on which the eye tracking system was running, one participant IBM T41 laptop on which the stimulus diagrams were shown, and adjustable chairs and tables. Adjustments were made to maintain the participant's eyes at a distance of approximately 50cm from the 14-inch monitor of the laptop. In order to reduce recording error that would be caused by head movement, a chin rest was used to stabilize the participant's viewing position and distance.

The eye tracking system used in the experiment was iViewX with Headmounted Eye-tracking Device (HED) (SensoMotoric Instruments GmbH (SMI)). The HED is a helmet to be worn by the participant that contains both an eye camera and a scene camera. Originally, the system displays the content of the scene camera in the scene video view of iViewX. However, since the participant's view in this study was confined to the laptop monitor, the system had been reconfigured to display the content of the laptop monitor screen by replacing the scene camera with a scan converter (Grand Hand View II, GrandTec).

A calibration tool called WinCal was used for visualizing calibration points and run on the participant laptop, so that the participant can calibrate while sitting in front of the laptop. The laptop and the operator PC were connected by a serial line for this purpose. Once enabled, WinCal can be triggered by the commands from the operator PC, maximize and minimize itself automatically at the start and end of calibration, respectively. The calibration area in iViewX had been set to match the resolution of the laptop monitor, 1024 ×768 pixels.

This eye tracker tracks eye movements by observing the position of the pupil and corneal reflex from the right eye. The combined video signal from the scan converter with eye position indicated by a gaze cursor was recorded into a MPEG video file for offline analysis.

## 2.3 Documents

The experimental documents consisted of a participation information sheet, a consent form, and tutorial sheets. Participants were presented with these materials to familiarize themselves with the tasks and the procedure.

The tutorial described the tasks, the procedure, and the online system that would be used by participants, explained the concepts, as well as presented six example drawings, in which several oral quizzes were given to make sure they understand testing questions properly.

## 2.4 Design

Graphs: Based on two social network graph data sets from the book of Wasserman and Faust (1994): Krackhardt's High-tech managers friendship relations, and Padgett's Florentine families business relations, we have got three graphs, among which the first is the full set of family business data with 11 nodes and 15 edges, the second is a subset of the same data with 9 nodes and 13 edges, and the third is a digraph and a subset of manager friendship data with 10 nodes and 14 edges.

Drawings: With our aims for this study on mind, the drawings were carefully designed. All drawings are shown in Table 1. In particular, two drawings on the same row have the same pair of nodes highlighted, with one drawing having crossings on a shortest path of the highlighted nodes and the other not, except drawings in the first row, where one drawing has a nearly straight shortest path and the other not. However, participants were not made aware of these.

Two popular layout conventions in social network visualization were also included, one is radial layout, where nodes are on different levels of circle with the important one in the centre (see drawings 1 and 12 in the first row of Table 1), and the other one is circle layout, where all nodes are on a single circle. Among all 12 drawings, drawings 1 and 12 were taken from Ka-Ping Yee et al (2001), and the others were drawn by hand; 8 drawings are about business relations, 4 others about friendship relations.

Questions: We designed the test questions according to the following criteria:

1. Simple enough so that participants can easily remember and understand.
2. Specific to the graph domain.
3. Representative.
4. Trigger not too many but not too few eye movements.

We finally decided that the question for family business relations is "What is the separation level between the two highlighted families?" which means the least number of links between the both; the question for manager friendship relations is "Do the two highlighted managers have Friend's friend relationship?" Two managers A and C have Friend's friend relationship if there is a manager B between them, i.e. either A→B→C or A←B←C.

Online system: Graph diagrams were displayed on the laptop monitor by a custom-built online experimental system. The system was designed so that:

1. For each drawing, the question is shown separately, not on the same screen, since people normally do not have their reading purposes shown with graph diagrams.
2. Keyboard usage is avoided, since people tend to look at the keyboard while typing, this causes error in terms of eye tracker. Instead, participants can answer questions by clicking buttons on the screen.
3. Each participant's response time for each drawing, which starts once a diagram is completely displayed, and ends once a button is

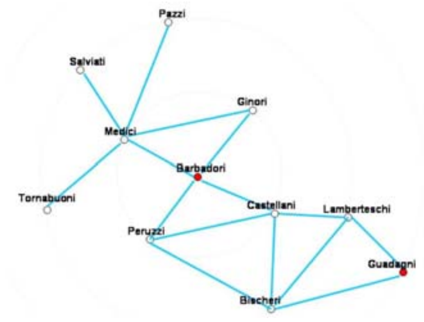
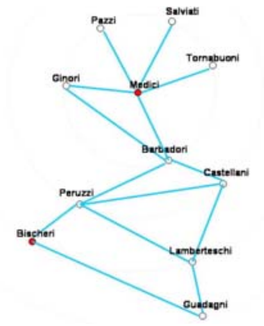
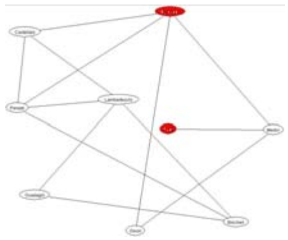
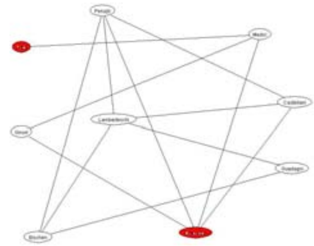
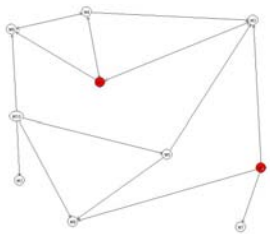
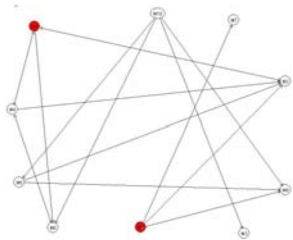
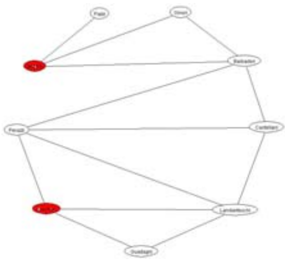
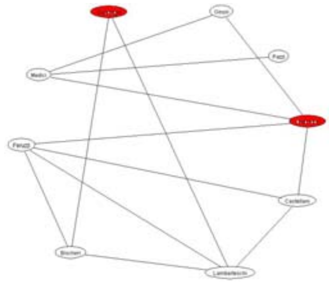
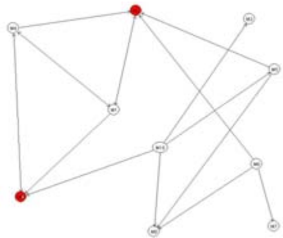
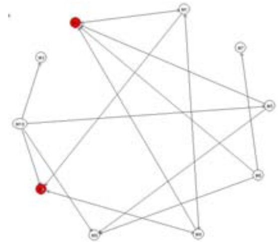
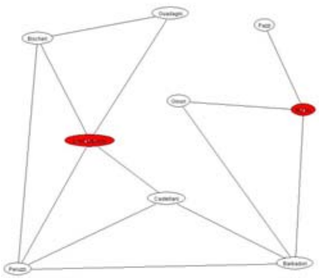
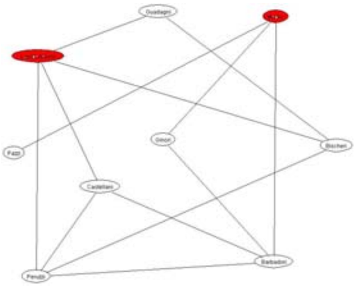
| | |
|---|---|
|   Drawing 1 |   Drawing 12 |
|  Drawing 2 |  Drawing 11 |
|  Drawing 7 |  Drawing3 |
|  Drawing 4 |  Drawing 8 |
|  Drawing 5 |  Drawing 9 |
|  Drawing 6 |  Drawing10 |

**Table 1: Drawings used in the experiment**

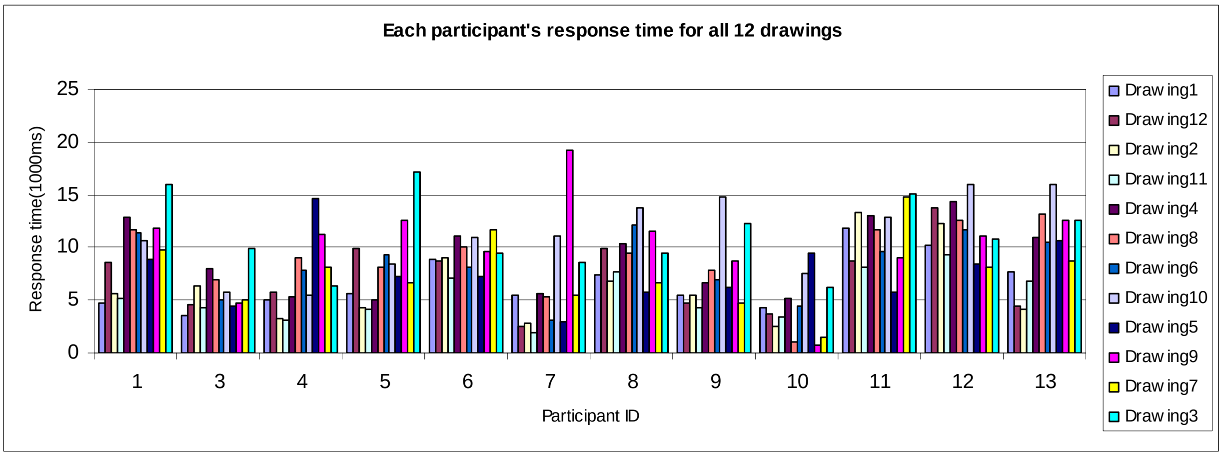



**Figure 1: Each participant's response time for all 12 drawings**

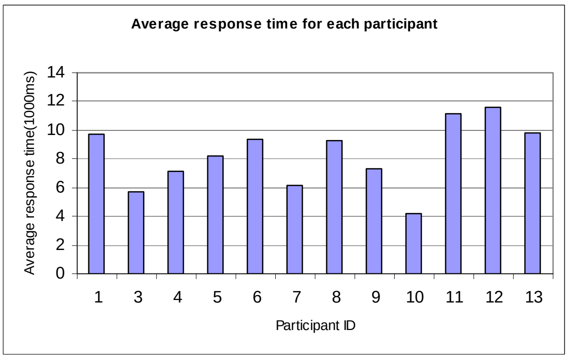



**Figure 2: Average response time for each participant**

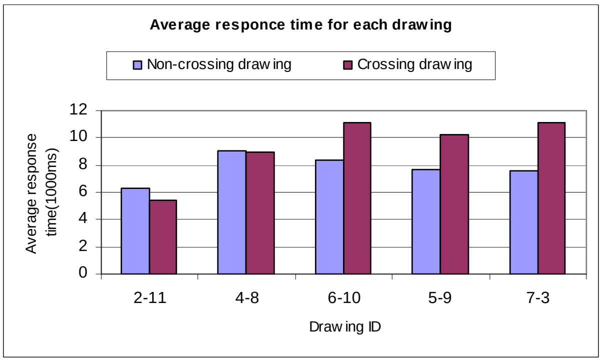



**Figure 3: Average response time for each drawing**

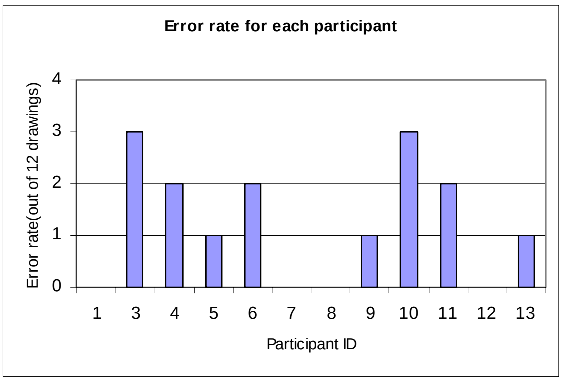



**Figure 4: Error rate for each participant**

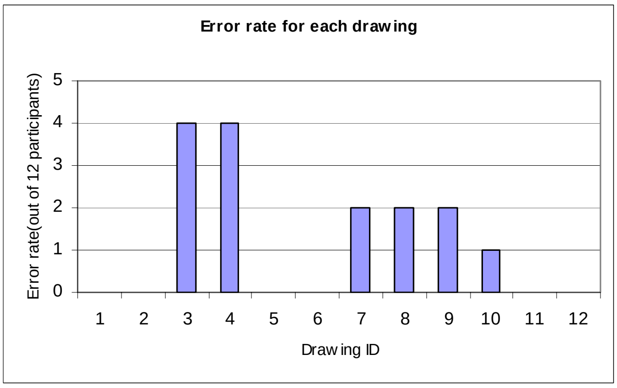



**Figure 5: Error rate for each drawing**

clicked, and corresponding response correctness are recorded.

The 12 drawings were assigned to four blocks, with three drawings in each and no single pair of crossing and non-crossing drawings within the same block. The arrangement of three drawings in one block was based on the observation that for each calibration, the eye movement data start getting worse after viewing three drawings based on our extensive pilot tests.

The drawings were shown in a random sequence. Just before proceeding with each block, a calibration was performed, and just after each block, there was a break, which was intended to reduce fatigue caused by the eye tracker, and refresh participants' memory so that they could not remember what they had done on previous blocks. The length of break and pace of reading tasks were in control of participants. Actually, the breaks took 2-3 minutes long.

During the preparation time, participants were instructed to look for the answer once a drawing is shown, click the corresponding button once the answer is determined, and not to look around in between.

## 2.5 Procedure

Before starting the experiment, the participants were asked to sign the consent form, read through and understand the materials, ask questions and do practice experiments (the drawings used for practice were different from the ones used in formal tests) as many times as they want so that they were sure that they understand the questions properly and feel comfortable with the testing environment.

Once ready to start, the participant was seated and the helmet was worn. After a short calibration, he/she started to run the online system and perform reading tasks. Every participant was asked to determine the answer to the question presented for each drawing and 12 drawings in total, and answer questions by clicking the corresponding button on the screen. The time taken for answering questions and the correctness were recorded.

Following graph reading tasks, a post-task questionnaire and experimental debriefing were given. The questionnaire asks participants what degree they are studying for; the status of enrolment; any previous experience of reading graphs; whether there are any particular drawing features that they think could aid or hinder their performance, and how these features can affect their reading behaviour.

The participants also were questioned about their reading behaviour they took during the experiment, and encouraged to verbalize any thoughts and feelings about the experiment. The notes had been taken and verified by the participants.

## 3 Results and Discussion

Thirteen students took part in the study, viewing 4 blocks of 3 drawings each, which means 4 × 13 = 52 video files had been recorded in total*. However, participant 2 read graphs with the aid of mouse cursor throughout the testing, which had made his testing settings different from others. We realized this only during the initial analysis phase and finally decided not to further consider his data for the sake of excluding all potentially unreliable data.

The summary results about their response time and correctness rate are shown in Figures 1 - 6. Since the two questions asked in this study were both about paths between two nodes (shortest path), for simplicity, we state our analysis in this more generic term.

Figure 1 shows the overall detailed performance for each participant, from which we can see that the response time varies for different participants and drawings.

Figure 2 suggests that experienced graph readers (participants 4, 5 and 7) had got shorter average response time. Participants 3 and 10 also had short response time, but with higher error rate, which can be seen in Figure 4.

Figure 3 seems to show that crossing in drawings has an impact on reader's performance. Table 2 also suggested that overall, a statistically significant difference does exist between the non-crossing drawings and crossing drawings ($t = -2.34$, $p = 0.021$). However, a closer inspection reveals a different story. In Figure 3, there is an exception: the average response time for drawing 2 was much longer than for drawing 11. This was a surprise to us because drawing 11 has many crossings on the shortest path, while drawing 2 has no crossing at all. This indicates that there must be some factor unknown at work.

| Drawings | Mean (ms) | Mean (ms) | T-Value | P-Value |
|---|---|---|---|---|
| 2 vs. 11 | 6305 | 5440 | 0.71 | 0.486 |
| 4 vs. 8 | 9026 | 8904 | 0.09 | 0.931 |
| 6 vs. 10 | 8350 | 11077 | -1.99 | 0.060 |
| 5 vs. 9 | 7627 | 10246 | -1.66 | 0.112 |
| **7 vs. 3** | **7602** | **11149** | **-2.47** | **0.022** |
| **Overall** | **7782** | **9363** | **-2.34** | **0.021** |

**Table 2: T-test analysis of the average response time for non-crossing and crossing drawings**

| Participant ID | RT (ms) for drawing 2 | RT (ms) for drawing 11 |
|---|---|---|
| 1 | 5558 | 5207 |
| 3 | 6319 | 4277 |
| 4 | 3274 | 3175 |
| 5 | 4246 | 4186 |
| 6 | 9063 | 7030 |
| 7 | 2794 | 1893 |
| **8** | **6760** | **7641** |
| 9 | 5448 | 4246 |
| **10** | **2564** | **3405** |
| 11 | 13309 | 8071 |
| 12 | 12208 | 9383 |
| **13** | **4116** | **6770** |

**Table 3: Participants' response time (RT) for drawings 2 and 11**

As shown in Table 3, only 3 (participants 8, 10, and 13) out of 12 participants had got shorter time for drawing 2 than for drawing 11, which suggests that the majority of participants' response were consistent with the average response result, and therefore this result is not the consequence of one or two participants spending significantly more time on drawing 2, or vice versa.

After examining the video files, we found that most participants started to search from the highlighted node on the top or left, and this is the case for drawings 2 and 11. We guess that this might have something to do with people's daily reading habit: starting from top to bottom, left to right. Unfortunately, for drawing 2, the top-left one is the node with 4 edges incident to it; they spent much time on following these edges trying to find the path to the other highlighted node. On the contrary, the top-left node for drawing 11 has only one edge, therefore they simply followed this edge to the next node which happens to be incident to the highlighted target node.

*: The video data can be accessed at http://www.it.usyd.edu.au/~whua5569/ex

However, drawing 11 does have many crossings on the shortest path. Surprisingly, a closer look of all the video data reveals that all the participants seemed to just ignore crossings on the way of paths while reading graphs. However, we noticed that almost all the crossings in our drawing set happened to be nearly 90 degree. According to Colin Ware et al (2002), this could explain why they had little impact on eye movements; at most, slow them down as claimed by most participants during the post-task interview.

Given this kind of crossings, another finding appears obvious. The edges and nodes alongside the paths can play an important role in hindering understanding. From the video data, we can see clearly that the participants tended to:

1. Follow edges which go toward the target node on the way and edges incident to the highlighted nodes.
2. Avoid dense area where there are relatively many edges and nodes.

In other words, the response time and extra eye movements depend on how much a particular layout can confuse readers. To be more specific, in terms of path searching tasks, the dense area alongside the paths, the edges going towards the target node and edges incident to the highlighted nodes can trigger extra eye movements and make a difference in people's reading performance.

Figures 4 and 5 show the error rate for each participant and each drawing, respectively. However, given the fact that all the participants are postgraduates, and we did not have time limit on accomplishing the tasks, there should not be any error. But this is not the case for this study. The only reasonable explanations are:

1. They did not take the test seriously.
2. They clicked a wrong button.
3. The particular layout hindered their understanding.
4. They did not feel well with experimental environment, just wanted to finish quickly.

Among the above, only the third one is what we looked for. However, we do not know whether the other three happened to the participants or not, so given the relatively small size (12) of samples, we can not only rely on the error rate data for analysis.

Figure 6 together with Figures 3 and 5 suggests that radial layout is helpful in aiding understanding. Also most participants claimed this kind of layout preferable. From Figure 6, it can be seen that on average, the participants spent a shorter time on drawing 1 which has a nearly straight shortest path than on drawing 12, although the t-test analysis indicates that there is no statistical difference between the two ($t = 0.71$, $p = 0.734$). This is consistent with the finding of Colin Ware et al (2002). Furthermore, from the video data for drawing 8, the fact that most participants failed to notice the shortest path via the leftmost node counter-proves the same.

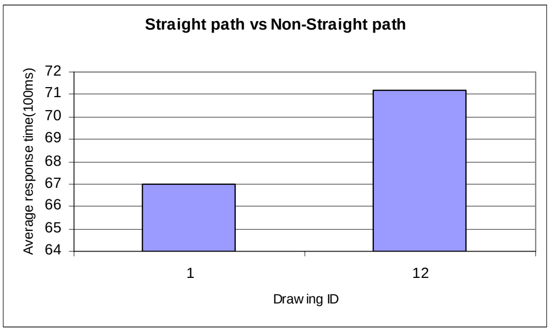



**Figure 6: Straight path vs. Non-Straight path**

Another interesting finding is that according to the eye movement data, it seems that most participants did not search for all possible paths, even not all possible shortest paths in order to determine which path is the shortest. Based on the notes taken during the post-task interview, there are two cases:

1. They did not notice other paths.
2. They noticed there could be another path, but did not think it is shorter so far. That means there is a reasoning and heuristic process during shortest path determination.

Based on the observations on the video data, the participants showed very similar high level search strategies:

1. Look for highlighted nodes; determine which one to start with.
2. Search for possible paths. During this process, some important edges and nodes might be repeatedly visited.
3. Determine and verify the answer.

However, for the step 2, participants 3, 4 and 10 showed quite different eye movement patterns. Their scan paths drifted a bit away from the corresponding edge paths, while others' eye movements follow the edge paths closely. They explained that they did not need to follow edges so closely since by looking at somewhere not far away, they still can see the paths. Obviously, this strategy made them spend a relatively short time on accomplishing tasks but with a higher error rate as shown in Figures 2 and 4, and this is not applicable for the friendship relation drawings, where they needed to follow edges to determine the arrow direction.

## 4 Limitations

Any formal empirical study has limitations, and can be improved upon, as no experiment can be perfect (Gottsdanker, R. 1978, Helen Purchase et al 1997). Therefore, in our case, the results we have presented need to be interpreted within the limitations which we identify below:

1. Questions: The two questions asked in our experiment were only about the relationship between two objects. However, the purpose of visualization, in general, is to support a range of

queries that are rapidly constructed on an ad hoc basis as part of visual thinking (Colin Ware et al 2004). For different query questions, people might use different reading strategies.

2. Graphs: The graphs used only had 11 nodes and 15 edges at most. However, in real world, especially for social network data, the scale of graphs can be too large to be shown on a single computer screen, to be comprehended without the aid of other tools, causing that the current experiment settings do not apply in this real world situation. This also raises another issue of whether the results obtained can be expected to scale with graph and query complexity.

3. Time limit: In this experiment, there was no time limit when participants performed their tasks. But this might not be the case in practice. For example, stock brokers need understand financial data in a short period of time.

4. Eye tracker: The eye tracker distracts participants and results in noisy data. Therefore a larger sample size and test runs are needed (Antti Aaltonen 1999).

## 5 Conclusion and Future work

To the best of our knowledge, this is the first eye tracking study on how people read relational node-edge graphs. In this study, we examined the eye movement data produced by participants, and found that a particular graph layout can affect the reading behavior in two ways: slow down and trigger extra eye movements. The reading performance depends on how much a particular layout can confuse the reader. In terms of path searching tasks, the edges incident to nodes concerned, edges going toward to the target node, and the edges alongside the paths affect drawing's readability and trigger extra eye movements. Although how crossings themselves affect eye movement patterns is not clear in this study, we should say that normally a graph layout becomes to be confusing when crossings are introduced.

In practice, people's graph reading behavior and understanding processes can be complex and differ in different environmental settings. Our work presented in this paper has only begun to examine certain aspects of the data produced within the limitations mentioned in the last section. Clearly, a great deal of work is yet to be done. In particular:

1. A specific drawing is there to support a variety of queries. In the study, we asked local questions about only two nodes. It is interesting to see whether there is a different eye movement pattern if a global question is asked (Raj M. et al 2003). By doing this, it is expected that we can obtain a more general understanding about how people interact with graphs.

2. In this paper, our analysis was mainly based on the video data. Further work focusing on quantitative analysis would be a welcome next step. However, this can be remarkably expensive. Firstly, eye tracking data contains a great deal of noise such as fixations between tasks, tracking loss of pupil, and therefore need to be filtered. Secondly, special software is needed to integrate the graph diagrams with corresponding scan paths.

3. We plan to perform an evaluation on how people interact with three dimensional graphs.

## 6 Acknowledgements

We are grateful to the School of Information Technologies and National ICT Australia for funding this research, to Dr. Seok-Hee Hong for her valuable advice on experiment design, to Dr. Nikola Nikolov for comments on the draft version of this paper, to Mr. Xiaoyan Fu for technical support, and to students who willingly took part in the experiment. Our special thanks go to Mr. Thorsten Dowe and Ms. Anne Kwik from SensoMotoric Instruments GmbH Germany for their professional support on the eye tracker reconfiguration. Ethical clearance for this study was granted by the University of Sydney, September 2004.

## 7 References

Aidan Feeney, Ala K., Simon P., John M. and Robert M. (2000): How people extract information from graphs: Evidence from a sentence-graph verification paradigm. *Lecture Notes in Artificial Intelligence*, 1889, 149-161.

Antti Aaltonen (1999): Eye tracking in usability testing: Is it worthwhile? http://www.cs.uta.fi/hci/mulmod/material/etusab.pdf. Accessed 20 Sept 2004.

Buswell, G. T. (1935): *How people look at pictures*. Chicago: University of Chicago Press.

Carpenter, P. A., and Shah, P. (1998): A model of the perceptual and conceptual processes in graph comprehension. *Journal of Experimental Psychology: Applied*, 4, 75-100.

Colin Ware, Helen Purchase, Linda Colpoys, and Matthew McGill (2002): Cognitive measurements of graph aesthetics. *Information Visualization*. Vol 1(2): 103-110.

Colin Ware, and Robert Bobrow (2004): Motion to support rapid interactive queries on node-link diagrams. *ACM Transactions on Applied Perception*. Vol 1(1): 3-18.

D. Ferrari and L. Mezzalir (1969): On drawing a graph with the minimum number of crossings. Technical Report 69-11, Istituto di Elettrotecnica ed Elettronica, Politecnico di Milano.

Gottsdanker, R. (1978): *Experimenting in psychology*. Prentice-Hall.

H. Trickey (1988): Drag: A graph drawing system. In *Proc. Int. Conf. On Electronic Publishing*, pages 171-182. Cambridge University Press.

Helen Purchase, Cohen R., and James M. (1997): An experimental study of the basis for graph drawing algorithms. *ACM Journal of Experimental Algorithm*, Vol 2(4).

Helen Purchase, Cohen R., and James M. (1995): Validating graph drawing aesthetics. *Proceedings of the Graph Drawing Symposium*, Passau, Germany, 435-446, Springer-Verlag.

Helen Purchase (1997): Which aesthetic has the greatest effect on human understanding? In *G. Di Battista, (Ed) Graph Drawing 97, 1353 of lecture Notes in Computer Science*, Springer Verlag, 284-290.

Jason S., Jeff B. and Mark D. (2003): Eye tracking observers during color image evaluation tasks. *Proceedings of SPIE, Vol 5007, Human Vision and Electronic Imaging VIII*, pp. 218-230.

Ka-ping Yee, Danyel F., Rachna D., and Marti H. (2001): Animated recentering for interactive navigation of radial graph layouts. Pre-publication draft, http://www.cs.berkeley.edu/~pingster/viz/ieee.html. Accessed 26 Sept 2004.

Peter Eades (1984): A heuristic for graph drawing. *Congressus Numerantium*, 42:149-160.

R. Lipton, S. North, and J.Sandberg (1985): A method for drawing graphs. In *proc. ACM Symp. On Computational Geometry*, pages 153-160.

R. Tamassia (1987): On embedding a graph in the grid with the minimum number of bends. *SIAM J. Computing*, 16(3):421-444.

Raj M., J. Gregory, and Deborah A. (2003): Thinking graphically: extracting local and global information. In *R. Alterman & D. Kirsch (Eds.), 25th Annual Meeting of the Cognitive Science Society*. Boston, MA: Erlbaum.

Wasserman, S. and Faust, K. (1994): *Social network analysis: methods and applications*. Cambridge University Press.

Yarbus, A. L. (1967): *Eye movements and vision*. New York: Plenum Press.