Evaluating Partially Drawn Links for Directed Graph Edges

Michael Burch, Corinna Vehlow, Natalia Konevtsova, and Daniel Weiskopf

VISUS, University of Stuttgart
Allmandring 19, 70569 Stuttgart, Germany

Abstract. We investigate the readability of node-link diagrams for directed graphs when using partially drawn links instead of showing each link explicitly in its full length. Providing the complete link information between related nodes in a graph can lead to visual clutter caused by many edge crossings. To reduce visual clutter, we draw only partial links. Then, the question arises if such diagrams are still readable, understandable, and interpretable. As a step toward answering this question, we conducted a controlled user experiment with 42 participants to uncover differences in accuracy and completion time for three different tasks: identifying the existence of a direct link, the existence of an indirect connection with one intermediate node, and the node with the largest number of outgoing edges. Furthermore, we compared tapered and traditional edge representations, three different graph sizes, and six different link lengths. In all configurations, the nodes of the graph were placed according to the force-directed layout by Fruchterman and Reingold. One result of this study is that the characteristics of completion times and error rates depend on the type of task. A general observation is that partially drawn links can lead to shorter task completion times, which occurs for nearly all graph sizes, tasks, and both tapered and traditional edge representations. In contrast, there is a tendency toward higher error rates for shorter links, which in fact is task-dependent.

1 Introduction

Visualizing graph data as node-link diagrams can lead to visual clutter [17]. This problem is most pronounced for dense graphs causing a huge number of edge crossings. Therefore, typical graph layout algorithms for node-link diagrams follow certain aesthetic criteria for graph drawing where the reduction of edge crossings is ranked very high. Other important aesthetic criteria include the minimization of edge lengths, the maximization of angles at link intersections, and the preservation of symmetries.

In our work, we ask the question if node-link graph visualizations are still useful, readable, and interpretable when reducing visual clutter by drawing links partially instead of showing each link explicitly in its full length and applying some sophisticated layout algorithm. To visually encode directed edges, we draw partial links beginning at the start vertex and pointing to the target vertex

M. van Kreveld and B. Speckmann (Eds.): GD 2011, LNCS 7034, pp. 226–237, 2012.

instead of ending exactly there. By doing this, many explicit link intersections are avoided, a fact that definitely reduces visual clutter but increases ambiguities. We speculate that this may lead to more graph misinterpretations but we are unsure how error rates and completion times behave when solving graph-related tasks. We also speculate that there is an optimal range of link lengths that balances the goals of reducing visual clutter on the one hand and minimizing error rates and completion times on the other hand.

We conducted a controlled user experiment with 42 participants to find out if the partial link visualization strategy for directed graphs has any benefits over the traditional complete link representations and, if so, what the best-suited link lengths would be. For the experiments, we used artificial data sets with constant characteristics: randomly generated graph data that follows the Barabási-Albert [1] model for scale-free networks where the degrees form a power-law distribution. The graphs were laid out with the force-directed algorithm by Fruchterman and Reingold [5], which meets relevant aesthetic criteria for graph drawing.

We used the following relevant independent variables in our study: edge style (tapered straight links according to Holten and Van Wijk [8] and straight links as used in traditional approaches [3]), varying number of vertices to reflect different sizes of graphs, and varying length of links to test for partially drawn links. We employed three different tasks in our study: (1) identifying the existence of a direct link between two highlighted nodes, (2) identifying the existence of an indirect connection with one intermediate node between two highlighted nodes, and (3) the detection of the node with the highest number of outgoing edges. The user study collected accuracies and completion times for those tasks to identify performance for the different settings of the independent variables.

2 Related Work

Graph visualization techniques aim at producing graph layouts that are readable, interpretable, and look aesthetically pleasing to the viewer; see Di Battista et al. [3] for an overview of graph visualization. We focus on the issue of visual clutter in graph layouts, which is getting more and more prominent with increasing data set size.

One approach to reducing clutter relies on partially drawn links. Early work in this direction is due to Becker et al. [2], who visualized graphs with half-links (called half-lines in their paper), i.e., a directed link is connected with its start vertex and points to its target vertex but cut at halfway. With line-shortening, they even used links with further reduced length. This visualization strategy reduces visual clutter by reducing the number of explicit link crossings; however, Becker et al. did not provide any user study to evaluate the effectiveness of their visualization approach. In recent work, Rusu et al. [18] investigated another variant of partially drawn links in diagrams of undirected graphs. They introduced short breaks in full links (instead of one piece of a short link in our case), relying on the Gestalt principle of closure to perceive the whole link; see

Koffka [12] for background information on Gestalt psychology. They provided a preliminary user-based evaluation; they conducted a subjective study, whereas we focus on a task-based evaluation with accuracies and completion times.

Graph drawing aesthetics also aim at reducing visual clutter for good readability. Rosenholtz et al. [17] developed a measurement technique for (generic) display clutter, based on color and luminance contrast features. They demonstrated that their measure can be used in an automated way to make design suggestions for drawing properties such as the location of an item. In general, the graph layout strongly affects the extent of clutter. There are many corresponding node-link graph layout algorithms; many of them employ a force-directed node placement, e.g., Eades' spring-embedder model [4], the Kamada-Kawai model [11], or the Fruchterman-Reingold model [5]. We base the graph layouts in our study on the Fruchterman-Reingold model because it aims at meeting several aesthetic criteria for graph drawing.

Purchase et al. [14,15,16] conducted several empirical studies on the aesthetics of graph layouts and discovered that the layout significantly affects user preferences and task performances. In their first study [16], they investigated effects of three common aesthetics criteria on the readability of graphs: symmetry, link crossings, and bends. They reported that minimization of bends and link crossing improves task performance, where the latter was identified as most important factor on graph reading performance [14].

Ware et al. [19] found out that not only edge crossing but also continuity is an important factor for aesthetic considerations. They indicated that clutter rather depends on the number of edges that cross a path itself, than the total number of edge crossings in the diagram. They reported that the angle between crossings affects readability. Their results were also supported by eye-tracking studies of Huang et al. [9,10] that showed that small angles cause slow eye movements. Holten et al. [7,8] performed several studies to evaluate the performance and preference of different directed edge representations. Their results showed a significant performance advantage for tapered and non-compressed animation representations compared to standard arrowheads. However, their study did not include the half-links of Becker et al. [2]. To close this gap, we focus our study on partially drawn straight links and less on the style of edge representation.

To evaluate the performance of graph layouts or edge representations, it is critical to choose adequate tasks. Lee et al. [13] suggested a list of low-level tasks and complex tasks to allow the generalization of experimental results. We picked three relevant tasks from their category of topology-based tasks for our study.

Finally, matrix visualization is another approach to graph visualization, substantially different from node-link diagrams. We restrict ourselves to evaluating node-link diagrams. For a user study on comparing matrix and node-link diagrams, we refer to Ghoniem et al. [6]. Amongst other results, they reported that tasks connected to finding paths were supported more effectively and efficiently in node-link diagrams than in matrix visualizations.

3 Graph Generation and Layout

We base our directed graph data used throughout the study on the Barabási-Albert graph model [1]. The graphs are laid out by the Fruchterman-Reingold algorithm [5]. We compare tapered [8] and traditional edge representations; and we vary the link lengths as well as the graph sizes.

3.1 Graph Model

Graphs are randomly generated by using the Barabási-Albert model [1], which produces scale-free networks following a power-law distribution for node fan-in and fan-out. By doing this, we guarantee that all graphs have similar statistical properties throughout the study. Based on this graph generation model, we implemented a Java program that generates directed graph data on demand.

3.2 Graph Layout

The graph data is then represented by applying the Fruchterman-Reingold algorithm [5], implemented in our Java-based study software. Figure 1 shows visualizations of example graphs generated by the Barabási-Albert model for three different graph sizes. Here, links are represented in tapered style that was also used by Holten et al. [7]. Please note that they restricted their study to links of full length.

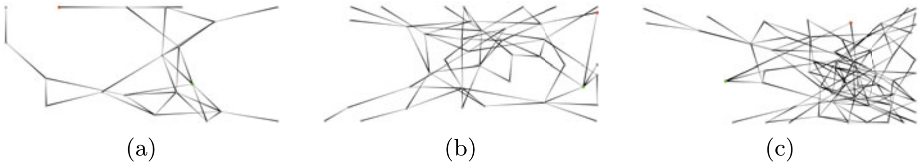

(a) (b) (c)

Fig. 1. Directed graph data produced by applying the Barabási-Albert model [1], laid out by the Fruchterman-Reingold algorithm [5], and displayed using tapered links of full length. Three different graph sizes are shown: (a) small graph with 20 nodes, (b) medium sized graph with 40 nodes, (c) large graph with 60 nodes.

3.3 Edge Representation

We use two different styles for representing directed graph edges:

- **Tapered Straight Links.** A needle-like shape that originates with its thicker end from the start vertex and points with its thinner end to the target vertex.
- **Traditional Straight Links.** An equally thick line that originates from the start vertex and heads to the target vertex.

Figure 2 compares tapered style and traditional style.

Fig. 2. Different edge styles: (a) traditional link, (b) tapered link. Both links are partial links with 75 percent of full link length.

3.4 Link Length

We vary the link lengths for both tapered and traditional edge representation styles. We use 100, 90, 75, 50, 25, and 12.5 percent of the length the link would have when drawn completely. For the traditional representation, we omit the 100% link length because otherwise the direction of a link could not be recognized by the viewer. Figures 3 (a)–(f) show examples of a graph consisting of 5 nodes and 10 links in tapered style; all 6 variations of link lengths are used. Figures 3 (g)–(k) show the same graph in the same layout, but with the traditional edge representation style; all variations of link lengths are used, except for the 100% link length.

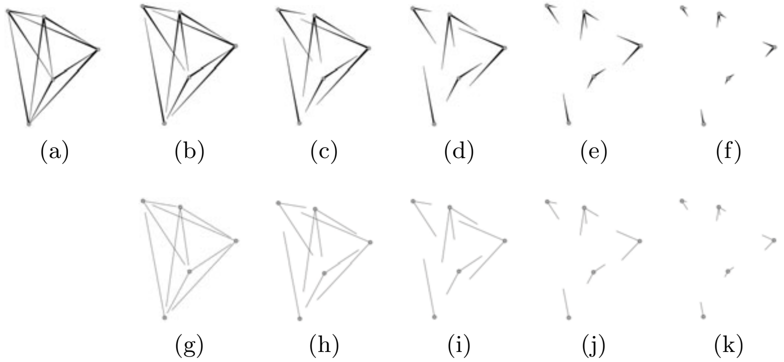


Fig. 3. Different link lengths are used in the study for both tapered and traditional edge styles: (a) tapered with 100% link length, (b) tapered 90%, (c) tapered 75%, (d) tapered 50%, (e) tapered 25%, (f) tapered 12.5%, (g) traditional 90%, (h) traditional 75%, (i) traditional 50%, (j) traditional 25%, and (k) traditional 12.5%.

3.5 Graph Size

Another variable that we vary is graph size, i.e., the number of vertices a graph contains. We choose 20 (small), 40 (medium), and 60 vertices (large) to test if there is any impact of graph size on accuracy and completion time. Figure 1 compares the three graph sizes for the example of full-length tapered edge representations. The graph density depends on the edges produced by the Barabási-Albert model.

4 User Experiment

We conducted a controlled user experiment with 42 participants to address the following research questions.

4.1 Research Questions

Since partially drawn links reduce visual clutter on the one hand, but increase the degree of ambiguity on the other hand, we considered the following research questions as relevant:

- **Research Question 1.** Can the tasks be answered more quickly with decreasing link length? In contrast, does the error rate increase due to more and more ambiguities for the target vertices?
- **Research Question 2.** Are the effects of Research Question 1 (i.e., decreasing completion time and increasing error rate) more pronounced for large graphs due to higher levels of overall visual clutter?
- **Research Question 3.** Are there any differences between tapered and traditional edge representations? Tapered links need more pixels to be drawn on screen. Hence, visual clutter is reduced more than in the traditional edge representation the shorter visible links are. Therefore, completion times should decrease more in the tapered style than in the traditional style. The error rates should stay similar since ambiguities occur equally in both styles.

4.2 Design

A repeated-measures design was used with three relevant independent within-subjects variables:

- **Edge Style.** Two possible edge representations: tapered straight links [8] and traditional straight links.
- **Number of Vertices.** Three graph sizes: 20 (small), 40 (medium), and 60 (large) vertices per graph.
- **Length of Links.** Six (five) different lengths of links, as percentage relative to the corresponding complete link: 100 (only for tapered), 90, 75, 50, 25, and 12.5 percent of the complete link.

We checked each of the three tasks in a separate block to reduce cognitive load from task switching. The three blocks were permuted to compensate learning and fatigue effects. Inside each task block, we randomized and balanced the graph sizes and link lengths, alternating between the two edge representation styles. This led to 3 [for graph sizes] $\times$ 3 [for tasks] $\times$ (5 [link length for traditional] + 6 [length for tapered]) = 99 configurations. Each of the 42 subjects performed each configuration twice (i.e., two repetitions), leading to 198 trials per subject and 8,316 trials in total.

We used a continue-on-demand study design, i.e., participants could decide when the next graph was represented by pressing a “Next” button. Participants were encouraged to take a longer break between the two repetition blocks.

4.3 Participants

We had 42 participants, 16 of whom were female and 26 male. The average age was 24.0 years; the youngest participant was at the age of 20 and the oldest at the age of 30 years. The participants were students of our university, except for one participant that had recently graduated. 16 of the participants were students of computer science or software engineering. All participants had normal or corrected-to-normal color vision, as confirmed by an Ishihara test and a Snellen chart; 15 of them wore glasses and 7 of them contact lenses. 7 participants claimed that they were familiar with graphs, 35 reported that were not (before the study). However, even the latter group was able to read node-link diagrams after a short introduction, as checked by asking graph-specific questions before the main test runs. Participants were compensated with EUR 10. Each experiment took between 44 to 100 minutes, depending on the speed of the participant. The average experiment time was 66 minutes.

4.4 Study Procedure

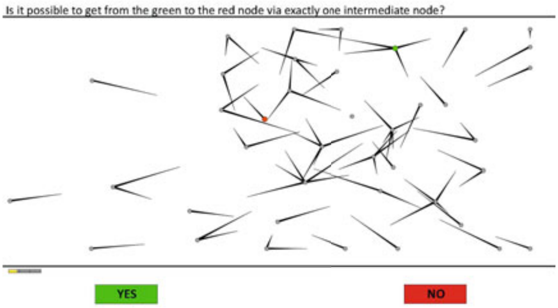
Participants were first asked to fill out a questionnaire about age, field of study, and prior knowledge in graph visualization techniques. Next, they read a short manual on the different graph diagrams, followed by test questions to check if they were able to read the node-link diagrams and solve the given tasks. Serving as a practice run-through, the initial test phase was conducted with a different set of stimuli data than the real experiment. Then, the actual experiment consisted of two larger blocks of trials (two repetitions as described in Section 4.2). During the experiment, subjects were sitting in front of a TFT screen with a resolution of 1920 × 1200 pixels at a distance of approximately 60 centimeters.

There was a “Give Up” option clearly present throughout the study; however, it was not used by the participants. There was no time limitation for the tasks. The participants were instructed to answer as accurately and as fast as possible. Once they found the solution, they had to confirm it by a mouse click to the correct position on screen (for Task 3, see Section 4.5) or by pressing a green-colored “YES” button or a red-colored “NO” button (for Tasks 1 and 2, see Section 4.5). The next stimulus was shown after the “NEXT” button had been pressed. Figure 4 shows a typical screenshot of the Java software employed for the user study.

4.5 Tasks

We tested three types of tasks in our study:

- **Task 1.** Is it possible to go from the node highlighted in green to the node highlighted in red by taking exactly one step, i.e., is there a directed edge starting at the green-colored node and pointing to the red-colored node?
- **Task 2.** Is it possible to go from the node highlighted in green to the node highlighted in red by taking exactly two steps, i.e., is there a path of length two starting at the green-colored node and ending at the red-colored node?



Is it possible to get from the green to the red node via exactly one intermediate node?

Fig. 4. Example screenshot from the user study. Here, a graph with 40 nodes and tapered links of 50% length is shown. The participant is asked for a path of length two starting at the green-colored node and ending at the red-colored node. (The figure is best viewed in the electronic color version of this paper.)

- **Task 3.** Which node has the highest number of outgoing edges?

Tasks 1 and 2 had to be answered by clicking on a button labeled with “YES” if the viewer agreed or “NO” if they disagreed. Task 3 had to be answered by clicking on the corresponding node on screen. We picked these tasks from the category of topology-based tasks for graphs according to Lee et al. [13]. Tasks 1 and 3 belong to the subcategory of adjacency-related tasks (direct connection between nodes); Task 2 is an accessibility-related task (here with an indirect connection).

5 Results

To evaluate the results, we first averaged the completion times and error rates over the 42 participants and 2 repetitions. This led to aggregated numbers for each of the 99 configurations. The scatterplot in Figure 5 and the line charts in Figure 6 show these averaged numbers.

In the scatterplot, the independent variables edge style, number of vertices, and length of links, as well as the task type are mapped to different visual attributes of glyphs: shape, border width, size, and color (shades of gray in black-and-white print). The scatterplot shows clusters for the different tasks: Task 1 (the direct edge search) has lowest completion times and error rates, with completion times ranging from 4.72 s (seconds) to 8.56 s, and error rates below 14%. The clusters for Task 2 and Task 3 spread more widely, but compared to Task 1, they show a clear tendency toward longer completion times and higher error rates with increasing number of vertices because thick-bordered glyphs lie right of and/or above thin-bordered glyphs.

For further analysis of the data, we turn to the line charts of Figure 6, which show completion times and error rates for all three tasks. Let us first focus on completion times (left column of Figure 6). Concerning the length of the links,

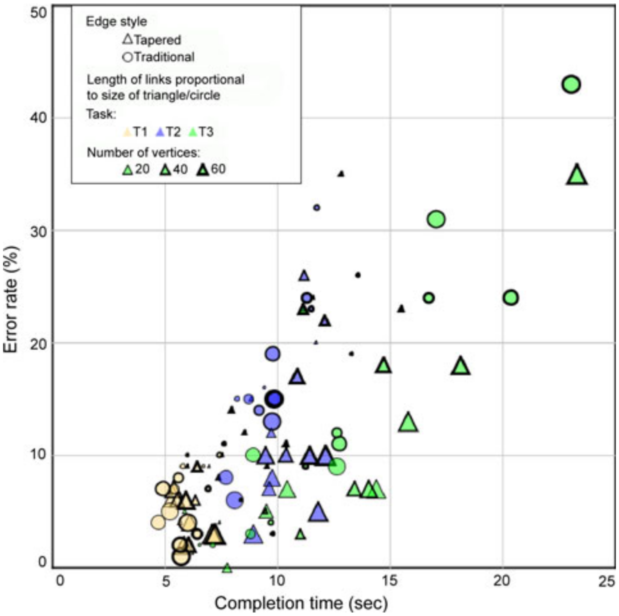


Fig. 5. Scatterplot of average error rates and completion times for all 99 configurations. The shape of the glyph denotes edge style, its size indicates length of the link, and its border width depends on the number of vertices. Different colors are used for the three different task types. (The figure is best viewed in the electronic color version of this paper; different colors correspond to shades of gray in black-and-white print.)

Tasks 1 and 2 show an interesting "dip" in the plots around 75% link length. This suggests that partially drawn links of 75% length provide the optimal balance between clutter reduction (supported by shorter links) and perception of node connections (supported by longer links) for these study parameters. Task 3 exhibits a different behavior: completion times become smaller and smaller with decreasing link length. This is reasonable because participants only had to find the "star" with the most jags to find the node with the highest number of outgoing edges. Therefore, we have indication that Research Question 1 (on link length) might be partially answered positively; however, there is a strong effect of type of task and, often, there might be an optimal length of medium size. Regarding edge style, the only visible difference appears for Task 3, where completion times are generally higher for traditional links. In contrast, Tasks 1 and 2 slightly tend toward lower completion times for traditional links. Therefore, the data to answer Research Question 3 (on edge style) is inconclusive; however, there might be an effect related to task type.

In general, completion times tend to increase with increasing number of vertices, independent of the type of task. The impact of graph size is most pronounced for Task 3, less so for Task 2, and even smaller for Task 1. Completion times are lowest for Task 1 (4.7 s – 8.5 s), medium for Task 2 (7.7 s – 15.5 s), and

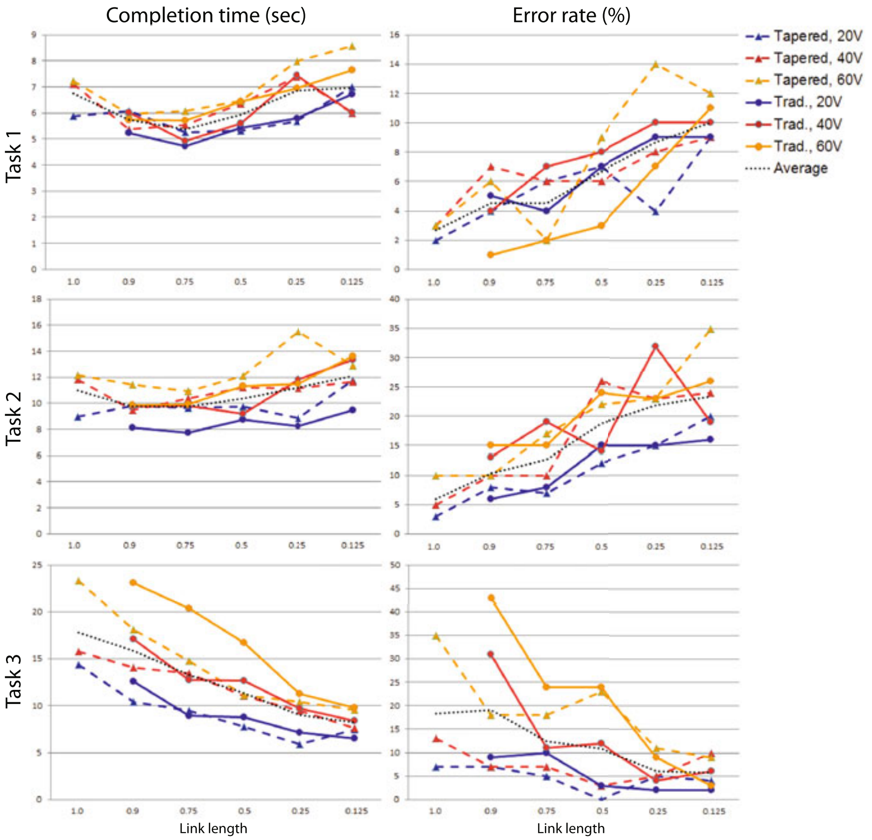


Fig. 6. The line charts plot the average completion times (left) and error rates (right) for all 99 configurations. Numbers were aggregated over 42 participants and 2 repetitions. The dotted line shows the average within the respective diagram. Link lengths are given as portions of the respective full length. Note that both completion times and error rates are in different scale for the three tasks.

longest for Task 3 (5.9 s – 23.3 s). We interpret this data as follows: Task 2 is more complex than Task 1, leading to longer completion times and larger spread thereof. This result is expected because Task 1 checks direct connections, whereas Task 2 indirect connections. Task 3 is even more complex, especially for large graphs, and thus leads to longer completion times and larger spread. However, overall we were not able to extract any structural impact of the number of vertices on how tasks were answered with different link length. Therefore, the data to answer Research Question 2 (on graph size) is inconclusive.

Let us now turn to error rates depicted in the right column of Figure 6. Error rates increase with decreasing link length for Tasks 1 and 2, whereas they improve for Task 3. Therefore, for Task 3, the error rates suggest a negative answer to Research Question 1. However, accuracy for Tasks 1 and 2 suggests a positive answer to Research Question 1. For Tasks 1 and 2, there is no clear difference in error rate for the two edge styles. In contrast, for Task 3, the traditional links lead to higher error rate for link lengths between 100% and 50%, but lower error rate for link lengths below 50%. Therefore, there is no clear answer to Research Question 3 in this case. Finally, the error rate depends on the size of the graph, as larger graphs lead to lower accuracy. Comparing the three tasks, the error rates are lowest for Task 1 ($1\% - 14\%$) and much higher for Task 2 ($3\% - 35\%$) and Task 3 ($0\% - 43\%$), which is consistent with the task-specific completion times. However, the error-rate data does not lead to a clear answer to Research Question 2.

6 Conclusion and Future Work

We have conducted a user study with 42 participants to test whether node-link diagrams are still readable and interpretable when drawing links only partially. One result of the study is that the influence of the link length on completion time and error rate clearly depends on the type of task. The study suggests that partially drawn links can lead to shorter completion times. Depending on the task, the optimal link length varies between still rather long links of 75% length (in Tasks 1 and 2) and much shorter links (as low as 12.5% length for Task 3). In general, however, accuracy tends to suffer when links are drawn only partially—except for Task 3, for which error rate improves for shorter links. We have also tested tapered and traditional edge styles, but found no relevant general effect of those. In conclusion, the main message is that there is potential usefulness of partially drawn links, especially when completion time is more important than accuracy; however, there are substantial task-dependent effects.

Therefore, we plan to include other tasks, especially more complicated tasks that may focus on cliques or clusters. Also, statistical hypothesis testing should complement our current qualitative discussion of results to come up with statistically significant evidence. Other venues of future work could include further variation of independent variables. For example, a larger range of graph sizes could be considered, models apart from the Barabási-Albert model of scale-free graphs could be used, or the graph density could be varied. Finally, graph layouts different from the Fruchterman-Reingold layout could be employed, e.g., a circular layout may be of interest because the graph vertices would be equidistantly placed on a circle circumference and, hence, target vertex ambiguities could be reduced.

Acknowledgements. The project was in part funded by the German Research Foundation (DFG) grant DFG WE 2836/4-1.

References

1. Barabási, A.L., Albert, R.: Emergence of scaling in random networks. Science 286(5439), 509–512 (1999)
2. Becker, R.A., Eick, S.G., Wilks, A.R.: Visualizing network data. IEEE Transactions on Visualization and Computer Graphics 1(1), 16–28 (1995)

3. Di Battista, G., Eades, P., Tamassia, R., Tollis, I.G.: Graph Drawing: Algorithms for the Visualization of Graphs. Prentice Hall, Upper Saddle River (1999)
4. Eades, P.: A heuristic for graph drawing. Congressus Numerantium 42, 149–160 (1984)
5. Fruchterman, T.M.J., Reingold, E.M.: Graph drawing by force-directed placement. Software: Practice and Experience 21(11), 1129–1164 (1991)
6. Ghoniem, M., Fekete, J.D., Castagliola, P.: A comparison of the readability of graphs using node-link and matrix-based representations. In: Proc. IEEE Symposium on Information Visualization, pp. 17–24 (2004)
7. Holten, D., Isenberg, P., van Wijk, J.J., Fekete, J.D.: An extended evaluation of the readability of tapered, animated, and textured directed-edge representations in node-link graphs. In: Proc. IEEE Pacific Visualization Symposium, pp. 195–202 (2011)
8. Holten, D., van Wijk, J.J.: A user study on visualizing directed edges in graphs. In: Proc. SIGCHI Conference on Human Factors in Computing Systems, pp. 2299–2308 (2009)
9. Huang, W., Eades, P.: How people read graphs. In: Proc. Asia-Pacific Symposium on Information Visualisation, pp. 51–58 (2005)
10. Huang, W., Hong, S.-H., Eades, P.: Layout Effects on Sociogram Perception. In: Healy, P., Nikolov, N.S. (eds.) GD 2005. LNCS, vol. 3843, pp. 262–273. Springer, Heidelberg (2006)
11. Kamada, T., Kawai, S.: An algorithm for drawing general undirected graphs. Information Processing Letters 31(1), 7–15 (1989)
12. Koffka, K.: Principles of Gestalt Psychology. Harcourt, Brace (1935)
13. Lee, B., Plaisant, C., Parr, C.S., Fekete, J.-D., Henry, N.: Task taxonomy for graph visualization. In: Proc. AVI Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization, BELIV 2006 (2006)
14. Purchase, H.C.: Which Aesthetic Has the Greatest Effect on Human Understanding? In: DiBattista, G. (ed.) GD 1997. LNCS, vol. 1353, pp. 248–261. Springer, Heidelberg (1997)
15. Purchase, H.C., Carrington, D., Allder, J.-A.: Empirical evaluation of aesthetics-based graph layout. Empirical Software Engineering 7(3), 233–255 (2002)
16. Purchase, H.C., Cohen, R.F., James, M.: Validating Graph Drawing Aesthetics. In: North, S.C. (ed.) GD 1996. LNCS, vol. 1190, pp. 435–446. Springer, Heidelberg (1997)
17. Rosenholtz, R., Li, Y., Mansfield, J., Jin, Z.: Feature congestion: a measure of display clutter. In: Proc. SIGCHI Conference on Human Factors in Computing Systems, pp. 761–770 (2005)
18. Rusu, A., Fabian, A.J., Jianu, R., Rusu, A.: Using the Gestalt principle of closure to alleviate the edge crossing problem in graph drawings. In: Proc. International Conference on Information Visualisation (IV 2011), pp. 488–493 (2011)
19. Ware, C., Purchase, H., Colpoys, L., McGill, M.: Cognitive measurements of graph aesthetics. Information Visualization 1(2), 103–110 (2002)