

Evaluating Overall Quality of Dynamic Network Visualizations

Weidong Huang¹(✉), Min Zhu², Mao Lin Huang^{3,4}, and Henry Been-Lirn Duh¹

¹ University of Tasmania, Hobart, Australia
{tony.huang,henry.duh}@utas.edu.au

² Sichuan University, Chengdu, China
zhumin@scu.edu.cn

³ Tianjin University, Tianjin, China
mao.huang@uts.edu.au

⁴ University of Technology Sydney, Ultimo, Australia

Abstract. Visualizing dynamic networks is a challenging task. One of the challenges we face is how to maintain visual complexity and overall quality of visualizations at a reasonable and sustainable level so that the information about the network embedded in the visualization can be effectively comprehended by the viewer. Many techniques and algorithms have been proposed and developed to facilitate the discovery of changing patterns. Much research has also been done in investigating how visualization should be constructed to be effective. However, how to measure and compare the quality of visualizations of a changing network at different time points has not been well researched. In this paper, we report on a preliminary work towards this direction. In particular, we apply an existing multi-dimensional overall quality measure in a user study data of different networks and found that the measured quality is positively correlated with user task performance regardless of network size.

Keywords: Cooperative visualization · Quality metrics · Evaluation · Visualization · Dynamic networks

1 Introduction

Cooperative visualization of dynamic networks, such as social networks and biological networks, is often seen in practice. For a given network and at any given time, some actors leave while new ones join the network. Accordingly, relationships among these actors also change. These dynamic features make visualization of these networks a challenging task [8]. Much research has been done on how dynamic networks can be better visualized while many tools are available to assist practitioners and researchers in making sense of this type of networks. In visualizing dynamic networks, one criterion that is commonly adopted is to preserve layout stability or mental map so that layout changes in a series of individual visualizations are kept to the minimum when possible [9]. Although empirical evidence for its positive effect is not conclusive [11], it is commonly

believed that keeping node positions unchanged across the visualizations will reduce the effort that is otherwise needed for a viewer to look for the same nodes in their new positions and refresh/rebuild their relationship patterns. This in turn will allow him/her to focus more on knowledge discovery and network comprehension while the network changes over a time period.

Although it may be an important factor for viewers to understand change patterns of a dynamic network quickly, maintaining mental map could be counter-productive as it can also have a negative effect on overall quality for individual visualizations. To explore this effect in the context of dynamic networks, as a first step towards this direction, we conducted a preliminary research into the method of measuring overall quality of networks of different sizes. More specifically, we took an existing quality measure of static networks [3] and apply it on an experimental data of dynamic networks. We examined whether the quality measure was correlated with the task performance.

In what follows, we first briefly review related work on quality metrics of network visualization, followed by the report on this preliminary research. The paper finishes with a short summary and our plan for future work.

2 Related Work

Dynamic networks are commonly visualized in the format of animation or a collection of static node-link diagrams [7]. For static graph drawings, aesthetics, such as minimum crossings and maximum crossing angles, are used to define layout for readability [5]. For dynamic networks, mental map has been widely accepted as an important criterion in judging the quality of layout. Much research has been done to develop visualization principles, tools and algorithms for both static and dynamic networks. For example, Friedrich and Eades [6] derived a set of criteria and measures for a good visualization of dynamic networks when animation is used. The criteria include uniform movement and symmetrical movement while measurements include minimize temporary edge crossings and maximize structured movements. Diehl et al. [12] developed an algorithms that is able to preserve mental map while the underlying graph structure changes. Animation also has implications on attention, memory and learning process which has been investigated in different domains [4, 10, 14].

While mental map requires that layout be changed as little as possible when the underlying graph is changed, aesthetics imply that different layouts should be used for different graphs. As a result, it is important to understand possible interactions between aesthetics and mental map in the context of dynamic networks. And measuring overall quality should help toward this goal. Different approaches have been proposed in the literature to measure visualization quality from different perspectives. Eades et al. [2] proposed a set of new shape-based quality metrics for large graphs, while Janicke and Chen [1] proposed a quality metric for visualization based on visual salience that can quickly guide the viewer's attention to the most relevant part of the image. Huang et al. [4] proposed an overall quality measure that takes into consideration performance gain in relation the cognitive cost devoted during the task performance.

Recently, Huang et al. [3] proposed an overall quality measure that aggregates z scores of individual quality metrics in a formula that outputs a single numeric value. The formula is as follows:

$$O = -z_{cross\#} + z_{crossRes} + z_{angularRes} - z_{uniEdge} \quad (1)$$

In this formula, $cross\#$ is the number of crossings in the visualization; $CrossRes$ is the minimum crossing angle. $AngularRes$ is the minimum angle formed by any two neighboring edges. $uniEdge$ is the standard deviation of all edge lengths. The authors also conducted a user study in which the measurement was tested on different visualizations of a same static network. The study demonstrated the sensibility and predictability of this measure with statistical significance, indicating its validity in measuring overall quality. In our study of exploring the possibility of measuring quality of different visualizations of a dynamic network, this formula is used. The details of the study are described in the following section.

3 Experiment

Visualizations of a dynamic network can be considered as visualizations of a series of static networks. Therefore, to examine the validity of formula 1 in measuring overall quality of dynamic networks, we applied it on visualizations, or drawings, of different networks (or graphs). We wanted to know whether this measure was still sensitive enough to predict task performance when different graphs were used. We used part of the experimental data of Huang et al. [5] for this purpose.

3.1 Data

There were 100 graphs, which were randomly selected from popular benchmark test suits: Rome graphs [13]. Rome graphs were a graph collection that were collected from real world applications in the field of software engineering, and have been widely used in graph drawing for testing purposes. The selected graphs were of different sizes, ranging from 15 to 50 ($Mean = 33.63$, $StDev = 6.47$). These graphs were drawn using a force-directed algorithm, resulting in 100 drawings in total.

The drawing stimuli were displayed one by one by a custom-built system. Forty-three subjects participated in the study. They were asked to find the shortest path between two pre-specified nodes in each drawing. Task completion time, responses to the task and mental effort were recorded. Visualization efficiency [4] was also computed afterward based on the recorded data. Therefore, the dependent variables for the study were time, effort, accuracy and efficiency. Overall quality, the predictor, was computed for each of these drawings using formula 1. Figure 1 shows three drawing examples that were of different graphs, but had close overall quality valued at 2.70, 2.74 and 2.80 from left to right,

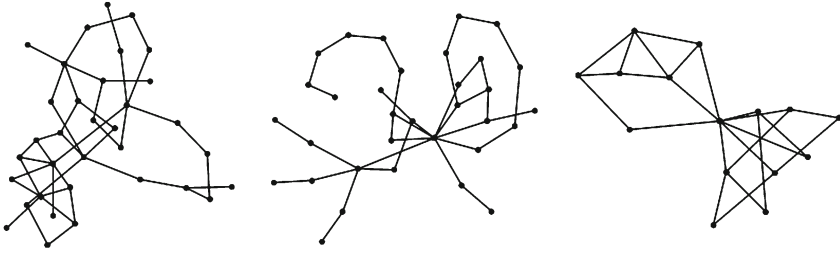


Fig. 1. Three drawings with close overall quality. The left drawing has 34 nodes and 42 edges, the middle drawing has 33 nodes and 37 edges, and the right drawing has 15 nodes and 24 edges

respectively. To test its predictability, we first looked at the scatter diagrams, and then regressed each dependent variable on overall quality. The results are reported in the next sub-section.

3.2 Results

The scatter diagrams are shown in Fig. 2. It can be seen that there was a general trend between each dependent variable and overall quality.

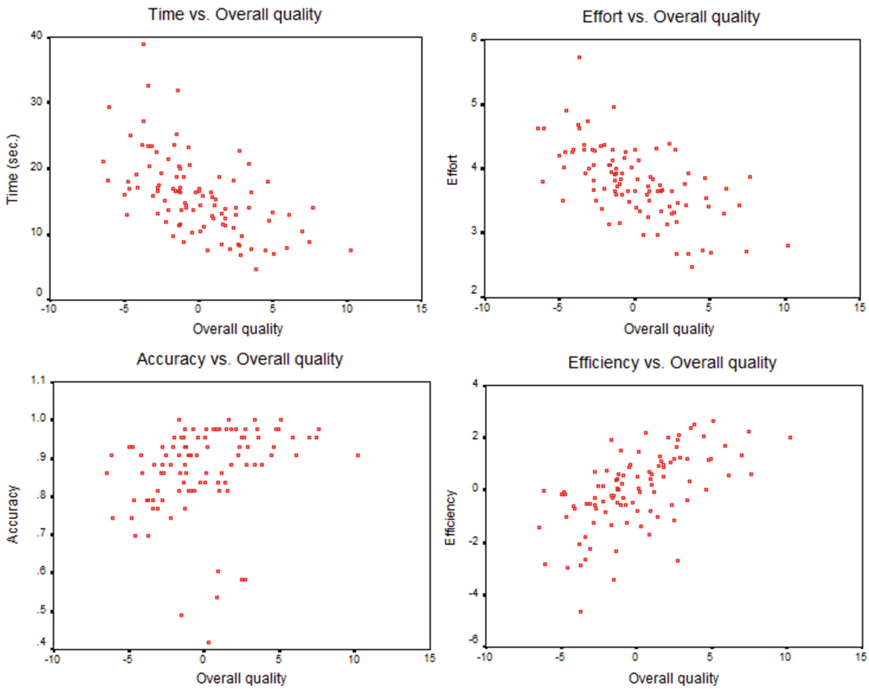


Fig. 2. Scatter diagrams between dependent variables and overall quality

Table 1. Results of simple linear regression tests

Dependent Var.	Predictor	β	F -statistics	R^2
Time	overall quality	-0.568	46.777 ***	0.323
Effort	overall quality	-0.610	58.072 ***	0.372
Accuracy	overall quality	0.279	8.242 **	0.078
Efficiency	overall quality	0.597	54.310 ***	0.356

Notes: **: $p < .01$; ***: $p < .001$

The results of simple linear regression tests are shown in Table 1. The overall regression test of time was significant, $F(1, 98) = 46.777$, $p < 0.001$. Time was negatively correlated with overall quality, $\beta = -0.568$. Overall quality accounted for 32.3% of the variance in time. Similarly, overall regression tests of effort, accuracy, efficiency were also significant with $p < 0.01$, or $p < 0.001$.

4 Discussion

The results of our data analysis indicated that the overall quality measure (formula 1) was significantly correlated with each of the dependent measures with a medium or large effect size. In other words, the overall quality measure has significant predictive capacity for the performance of human graph comprehension with graphs of different sizes.

This finding indicates that the measure can be used to compare the overall quality of different visualizations, regardless of the size of the underlying networks and their internal structures. Being able to compare overall quality without having to hold the underlying graph constant makes the measure more widely applicable. For example, given a large graph dataset, interaction techniques are often used to show one part of the graph at one time to facilitate the human graph reading process. However, designing and implementing an interaction method can be time-consuming. Even if it has been implemented, there is still a need of knowing whether the interaction method in consideration is actually effective or not by conducting more time-consuming user studies. By applying the overall quality measure on the layout of the whole graph and the intended layouts of its sub-graphs, the visualization designer can quickly decide whether the proposed interaction method is effective, without having to go through the whole implementation process as required otherwise.

Further, in visualizing dynamic networks, it is often desirable to keep positions of nodes unchanged in order to preserve mental map. However, empirical research on dynamic visualization has so far not been conclusive in supporting this practice [11]. Being able to compare the overall quality of snapshots of an evolving graph, the measure of formula 1 may offer us new insights into visualizing dynamic networks and explaining the current research results on mental map.

It should be noted that our study had limitations. For example, only one specific path-search task was used and further, this task might not be generally related to comprehension of dynamic networks in practice.

5 Summary

In this paper we presented an initial study that is part of our effort of investigating how preserving mental map and at the same time maintaining overall quality at a reasonable level can be balanced for visualization of dynamic networks. We applied an existing overall quality metric that is an aggregation based measure on a number of visualizations of different networks. The results showed that the quality measure was significantly correlated with the user task performance, indicating its validity for the purpose of measurement. For future work, more comprehensive studies with various types of benchmark datasets, use cases and tasks should be conducted to test the validity of the measurement.

References

1. Janicke, H., Chen, M.: A salience-based quality metric for visualization. In: Proceedings of the 12th Eurographics/IEEE - VGTC Conference on Visualization (EuroVis 2010), pp. 1183–1192 (2010)
2. Eades, P., Hong, S.-H., Klein, K., Nguyen, A.: Shape-based quality metrics for large graph visualization. In: Di Giacomo, E., Lubiw, A. (eds.) GD 2015. LNCS, vol. 9411, pp. 502–514. Springer, Heidelberg (2015). doi:[10.1007/978-3-319-27261-0_41](https://doi.org/10.1007/978-3-319-27261-0_41)
3. Huang, W., Huang, M.L., Lin, C.-C.: Evaluating overall quality of graph visualizations based on aesthetics aggregation. *Inf. Sci.* **330**, 444–454 (2016)
4. Huang, W., Eades, P., Hong, S.-H.: Measuring effectiveness of graph visualizations: a cognitive load perspective. *Inf. Vis.* **8**(3), 139–152 (2009)
5. Huang, W., Eades, P., Hong, S.-H., Lin, C.-C.: Improving multiple aesthetics produces better graph drawings. *J. Vis. Lang. Comput.* **24**(4), 262–272 (2013)
6. Friedrich, C., Eades, P.: Graph drawing in motion. *JGAA* **6**(3), 353–370 (2002)
7. Moody, J., McFarland, D., Bender-DeMoll, S.: Dynamic network visualization. *Am. J. Sociol.* **110**, 1206–1241 (2005)
8. Bender-deMoll, S., McFarland, D.: The art and science of dynamic network visualization. *J. Soc. Struct.* **7**(2) (2006)
9. Misue, K., Eades, P., Lai, W., Sugiyama, K.: Layout adjustment and the mental map. *J. Vis. Lang. Comput.* **6**(2), 183–210 (1995)
10. Minamoto, T., Shipstead, Z., Osaka, N., Engle, R.: Low cognitive load strengthens distractor interference while high load attenuates when cognitive load and distractor possess similar visual characteristics. *Attention Percept. Psychophys.* **77**(5), 1659–1673 (2015)
11. Archambault, D., Purchase, H.: The map in the mental map: experimental results in dynamic graph drawing. *Int. J. Hum. Comput. Stud.* **71**(11), 1044–1055 (2013)
12. Diehl, S., Gorg, C., Kerren, A.: Preserving the mental map using foresighted layout. In: *VisSym*, pp. 175–184 (2001)
13. di Battista, G., Garg, A., Liotta, G., Tamassia, R., Tassinari, E., Vargiu, F.: An experimental comparison of four graph drawing algorithms. *Comput. Geom. Theory Appl.* **7**(5–6), 303–325 (1997)
14. Ng, H.K., Kalyuga, S., Sweller, J.: Reducing transience during animation: a cognitive load perspective. *Educ. Psychol.* **33**(7), 755–772 (2013)