Validation of Quantitative Measures for Edge Bundling by Comparing with Human Feeling

Ryosuke Saga

Osaka Prefecture University, Japan

Abstract

This paper describes an analysis of the relationship between human cognition and quantitative measures for a visualization method called edge bundling.Aesthetic rules-based measures, namely, mean edge length difference (MELD), normalized MELD (NMELD), mean occupation area, and edge density distribution (EDD), for evaluating and quantifying the result of edge bundling are proposed. However, comparing these measures with human cognition has not been analyzed. Therefore, a questionnaire survey with approximately 40 respondents was conducted to clarify the relationship between human cognition and these evaluation measures. Results showed that NMELD, MELD, and EDD demonstrate robust and significant correlations with human cognition.

CCS Concepts

•***Human-centered computing*** → ***Visualization design and evaluation methods;*** *Graph drawings;*

1. Introduction

Network visualization is one of the visualization methods that can express data and the relationships between data. Network visualization is based on graphical representation in mathematics. Hence, a network consists of vertices and edges, which also exhibit attributes. The vertex set is V, the edge set is E, and a network G is presented as G= (V,E).

Several methods can be used to visualize networks [HMM00]. Edge bundling has drawn recent attention as a novel approach to improve visual clutter. This method enables observers to recognize the main stream of edges through bundle edges in accordance with certain standards. For example, several methods based on the hierarchical structure of nodes [HvW09], parallel coordinates [ZYQ*08], mechanical models [TE10], spring force-based approach [HvW09] have been proposed.

These edge bundling methods are considered useful in certain cases. However, the efficiency of each method is difficult to compare. Time (e.g., calculation time and complexity), comprehension (such as readability), and presentation (e.g., expression, representation, and visual encoding) are generally considered in evaluation. Time can be evaluated logically and calculatedly by a computer, but comprehension and presentation need to be qualitatively evaluated by a questionnaire and interview. For example, McGree and Dingliana performed user experiments to evaluates the impact of edge bundling [MD12]. The authors conducted experiments to validate certain hypotheses on path tracing and cluster connectivity. Although this work shows the interesting result that edge bundling negatively impacts user performances, this work must collect the questionnaires. This condition indicates that a considerable time and cost is required to evaluate the edge bundling methods.

From the background, to evaluate proposed edge bundling methods quantitatively and effect, Saga has proposed three measures, including mean edge length difference (MELD), mean occupation area (MOA), and edge density distribution (EDD) [Sag16]. These measures are based on aesthetic rules [PCA00], however, the relationships between human cognition and these three measures are yet to be validated. If a correlation is established between the qualitative evaluations of humans and these measures, then we can apply these three measures instead of relying on our own evaluation.

2. Quantitative Evaluation Measure of Edge Bundling

Edge bundling methods are quantitatively evaluated by using three measures, namely, MELD, MOA, and EDD. MELD reveals the differences in edge lengths before and after edge bundling. A slight difference in edge length indicates superior edge bundling results because of over-bundling, while a huge difference implies that the meaning of the original network has been lost. That is, The less MELD is, the better an edge bundling method is. MOA reveals the differences among the compressed areas after edge bundling. An excellent bundling procedure can compress those areas that are occupied by edges. Therefore, the less MOA is, the better an edge bundling method is. EDD is rooted on the ideas that an excellent edge bundling method can gather the edges within a unit area and that the density per unit is high. Therefore, the larger EDD is, the better an edge bundling method is. These three measures are calcu-

DOI: 10.2312/eurp.20181121

lated as follows:

$$MELD = \frac{1}{n}\sum_{e \in E}|L'(e) - L(e)| \tag{1}$$

$$MOA = \frac{1}{N}|\bigcup_{e \in E} O(e)| \tag{2}$$

$$EDD = \frac{1}{n}\sum_{a \in A}|p(a) - p| \tag{3}$$

In Equation (1), n is the number of edges, $L(e)$ is the length of edge e before edge bundling, and $L(e)$ is the length of edge e after edge bundling. In Equation (2), N is the number of areas, $O(e)$ is the set of areas occupied by edge e over an occupation degree which value is 5% of unit area in this application), and | | indicates the number of elements contained in a set. In Equation (3), A is a set of unit areas and $p(a)$ is the rate of the number of pixels, in which the edges pass in Area a, and p is a mean of $p(a)$. In Equations (2) and (3), the unit size is set to 6, that is, each unit area has 6ÃŮ6 pixels. This paper also proposes a new measure called normalized MELD, which normalizes the difference between $L(e)$ and $L(e)$ and prevents MELD from being affected by long links.

$$NMELD = \frac{1}{n}\sum_{e \in E}\frac{L'(e) - L(e)|}{L(e)} \tag{4}$$

3. Experiment

3.1. Experiment goal and its process

A questionnaire survey was conducted to confirm the relationships between human cognition and the aforementioned measures. The questionnaire contained 10 questions about the three graph drawing results shown in Fig. 1. Each question shows the three graph including original layout and other two results of edge bundling methods. In this Fig. 1, the first graph from the left shows the original layout, while the second and third graphs show the Force-Directed Edge Bundling(FDEB) and Cluster-based Edge Bundling (CBEB) [YS17] results, respectively. The respondents were asked to rank these drawings from best (score of 1) to poorest (score of 3). The

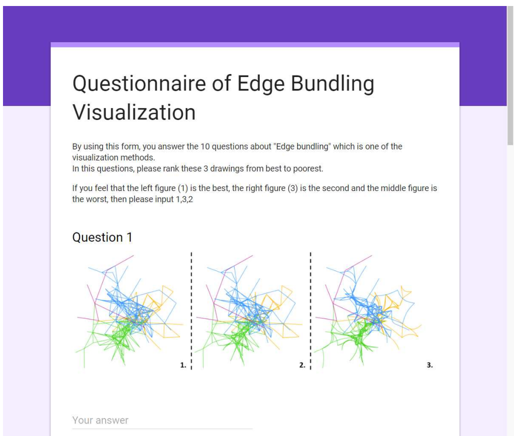

Figure 1: *An example of a question in questionnaire*

Table 1: *The number of persons feeling "the best drawing"*

Question	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
Original	12	13	12	15	11	17	22	15	15	14
FDEB	22	17	17	15	20	13	7	14	15	12
CBEB	5	9	10	9	8	9	10	10	9	13

Table 2: *Correlation Analysis Result*

	NMELD	MELD	MOA	EDD
Correlation	-0.697	-0.636	0.234	-0.644
t-value	-2.569	-2.178	0.638	-2.229
p-value	0.033**	0.061*	0.541	0.056*

* p <0.1, ** $p < 0.05$

average ranking of each of the four measures and the differences between FDEB and CBEB were calculated for every question. We also calculated the differences among NMELD, MELD, MOA, and EDD in terms of their FDEB and CBEB and then checked the correlations between the rankings provided by the respondents and those presented by the aforementioned measures. Scoring methods, such as the Likert scale, was considered inapplicable because accurately evaluating the scores is difficult for users. This analysis denotes that designing the questionnaires to make the intervals for each score (e.g., the interval between five and four stars) equal is difficult. However, ordinal scale, such as ranking, can be evaluated by comparing them, and ordinal scaling is used in several studies to understand user preference accurately [KAM03]. Therefore, we used a ranking score rather than a scoring method, such as the Likert scale, in these questionnaires.

Questionnaires were distributed among the students who don't know edge bundling well in the affiliation by using Google Forms, and a total of 39 completed questionnaires were obtained.

3.2. Analysis Result

We checked the number of persons that ranked the best drawing for each question at first. Table 1 shows the results of questionnaires and Tables 2 shows the result of correlation analysis between the two values; the differences of average ranking and the differences of the values given by four measures (NMELD,MELD, MOA, and EDD) between FDEB and CBEB. The results reveal some correlations between the rankings given by the respondents and those given by the three measures. From the results, we can consider that NMELD, MELD, and EDD have a certain correlation with human cognition. Also, the p-values for NMELD, MELD, and EDD are less than 5% or 10%, thereby indicating a significant correlation between these three measures and human cognition at the 5% or 10% level. However, MOA has a very large p-value, thereby indicating an insignificant relationship between MOA and human cognition.

In summary, the results indicated that we can use the NMELD, MELD, and EDD for quantitative evaluations with a certain level of reliability. As future work, a questionnaire survey on graph layouts and investigation of the correlations between these layouts and human cognition using more students than the respondents in this study may be conducted to acquire more reliable results. Also, we have to conduct other designed questionnaires to analyze from different viewpoints.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number 16K01250.

References

[HMM00] HERMAN I., MELANÇON G., MARSHALL M. S.: Graph visualization and navigation in information visualization: A survey. *IEEE Trans. Vis. Comput. Graph. 6*, 1 (2000), 24–43. URL: http://dx.doi.org/10.1109/2945.841119, doi:10.1109/2945.841119. 1

[HvW09] HOLTEN D., VAN WIJK J. J.: Force-directed edge bundling for graph visualization. In *Proceedings of the 11th Eurographics / IEEE - VGTC Conference on Visualization* (Chichester, UK, 2009), EuroVis'09, The Eurographs Association & John Wiley & Sons, Ltd., pp. 983–998. URL: http://dx.doi.org/10.1111/j.1467-8659.2009.01450.x, doi:10.1111/j.1467-8659.2009.01450.x. 1

[KAM03] KAMISHIMA T.: Nantonac collaborative filtering : Recommendation based on order responses. *Proc. of The 9th Int'l Conf. on Knowledge Discovery and Data Mining* (2003). URL: https://ci.nii.ac.jp/naid/10011699409/. 2

[MD12] MCGEE F., DINGLIANA J.: An empirical study on the impact of edge bundling on user comprehension of graphs. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (New York, NY, USA, 2012), AVI '12, ACM, pp. 620–627. URL: http://doi.acm.org/10.1145/2254556.2254670, doi:10.1145/2254556.2254670. 1

[PCA00] PURCHASE H. C., CARRINGTON D. A., ALLDER J.: Experimenting with aesthetics-based graph layout. In *Theory and Application of Diagrams, First International Conference, Diagrams 2000, Edinburgh, Scotland, UK, September 1-3, 2000, Proceedings* (2000), pp. 498–501. URL: https://doi.org/10.1007/3-540-44590-0_46, doi:10.1007/3-540-44590-0_46. 1

[Sag16] SAGA R.: Quantitative Evaluation for Edge Bundling Based on Structural Aesthetics. In *EuroVis 2016 - Posters* (2016), Isenberg T., Sadlo F., (Eds.), The Eurographics Association. doi:10.2312/eurp.20161131. 1

[TE10] TELEA A., ERSOY O.: Image-based edge bundles: Simplified visualization of large graphs. In *Proceedings of the 12th Eurographics / IEEE - VGTC Conference on Visualization* (Chichester, UK, 2010), EuroVis'10, The Eurographs Association & John Wiley & Sons, Ltd., pp. 843–852. URL: http://dx.doi.org/10.1111/j.1467-8659.2009.01680.x, doi:10.1111/j.1467-8659.2009.01680.x. 1

[YS17] YAMASHITA T., SAGA R.: Cluster-based edge bundling based on a line graph. In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017) - Volume 3: IVAPP, Porto, Portugal, February 27 - March 1, 2017.* (2017), pp. 311–316. URL: https://doi.org/10.5220/0006236003110316, doi:10.5220/0006236003110316. 2

[ZYQ*08] ZHOU H., YUAN X., QU H., CUI W., CHEN B.: Visual Clustering in Parallel Coordinates. *Computer Graphics Forum* (2008). doi:10.1111/j.1467-8659.2008.01241.x. 1