Information Sciences 330 (2016) 444–454

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Evaluating overall quality of graph visualizations based on aesthetics aggregation

Weidong Huang [a,*], Mao Lin Huang [b,c], Chun-Cheng Lin [d]

[a] School of Engineering and ICT, University of Tasmania, Australia
[b] School of Computer Software, Tianjin University, China
[c] School of Software, FEIT, University of Technology, Sydney, Australia
[d] Department of Industrial Engineering and Management, National Chiao Tung University, Taiwan

ARTICLE INFO

Article history:
Received 20 November 2014
Revised 5 May 2015
Accepted 16 May 2015
Available online 21 May 2015

Keywords:
Graph drawing
Overall quality
Aesthetics
Measurement
Effectiveness

ABSTRACT

Aesthetics are often used to measure the layout quality of graph drawings and it is commonly accepted that drawings with good layout are effective in conveying the embedded data information to end users. However, existing aesthetic criteria are useful only in judging the extents to which a drawing conforms to specific drawing rules. They have limitations in evaluating overall quality. Currently graph visualizations are mainly evaluated based on personal judgments and user studies for their overall quality. Personal judgments are not reliable while user studies can be costly to run. Therefore, there is a need for a direct measure of overall quality. In an attempt to meet this need, we propose a measurement that measures overall quality based on individual aesthetics and gives a single numerical score. We present a user study that validates this measure by demonstrating its sensibility in detecting quality changes and its capacity in predicting the performance of human graph comprehension. The implications of our proposed measure for future research are discussed.

1. Introduction

In order to take advantage of the powerful human visual perception system, node-link diagrams are often used as a visual tool for the purposes of communication and understanding of non-visual graph data. When used for graph data, node-link diagrams are also called graph drawings, and are sometimes simply called graphs, drawings or visualizations if no confusions are caused. However, drawing graphs into node-link diagrams does not automatically make the process of communication and understanding better as a graph can be laid out in very different ways. Empirical research has shown that layout affects how a graph is perceived [27]: a good layout facilitates the process, while a poor layout may hinder the process. Therefore, it is important to know how effective an intended layout is in conveying the underlying data to end users when drawing a graph.

Layout rules, or aesthetics, have been used as quality criteria to guide the choice between layouts. It is commonly accepted that drawings conforming to these aesthetics are of good quality and can be effective [3]. Two examples of these aesthetics are the minimum number of edge crossings and the maximum display of symmetries. In other words, a drawing with fewer crossings and more symmetries are better. However, the existing aesthetics are useful only in judging the extents

* Corresponding author.
E-mail address: tony.huang@utas.edu.au (W. Huang).

http://dx.doi.org/10.1016/j.ins.2015.05.028
0020-0255/

to which a drawing conforms to specific drawing rules; they have limitations in evaluating overall quality. One of the causes for the limitation was the fact that most of the aesthetics conflict with each other; it is not possible to implement all of them to the fullest at the same time. Optimizing one aesthetic can be achieved only at the cost of other aesthetics, leaving the overall quality uncertain. Fig. 1 gives a simple example of conflicting aesthetics. It shows two drawings of a graph. To draw the graph with maximum symmetries, more crossings are required (left). However, maximum symmetries are no longer possible when it is drawn with minimum crossings (right). The conflicting relationship affects current practices of graph visualization greatly. On the one hand, many algorithms for automatic graph drawing are designed to optimize only one or two aesthetics, and different algorithms focus on different aesthetics. This makes it difficult for an algorithm user to choose which algorithm to use when he or she has more than one algorithm candidate at hand. The reason for this is because there is no easy way available to tell whether, for example, an algorithm that is to minimize the number of crossings will produce better drawings than another algorithm that is to maximize symmetries in terms of overall quality.

On the other hand, it is generally acknowledged that the best layout is the balance of aesthetics. This is partly reflected in the fact that force-directed algorithms have been the most widely used tools for graph visualization. This particular type of algorithms simulate a graph as a physical system and assign numerical weight values to forces that represent the aesthetics considered. These forces work together and a balanced layout is reached in the end. This final layout is a compromise between the forces, or the aesthetics in consideration. In addition, given several drawings of different layouts, by computing and comparing values of aesthetics, we are able to find out which drawing has achieved a better trade-off between aesthetics considered [11]. However, seeking a compromise between aesthetics only gives us a better chance of producing good drawings [21]. Different aesthetics affect human graph comprehension to different degrees. Without conducting a user study, we are unable to know, for example, whether a drawing produced based on one set of aesthetics is better than another drawing that is produced based on a different set of aesthetics. There are no empirically verified guidelines or quality measures available for us to make such type of evaluation at the design stage.

Due to the lack of appropriate measures or methods, graph visualizations are evaluated mainly based on personal judgments and user studies. However, personal judgments are subjective and are not reliable, while user studies can be costly to run and can only be done after the visualization has been completed. Therefore, there is a need for a reliable and objective measure so that we can evaluate overall quality at the early design stage of a visualization process. This measure will help visualization designers to quickly judge or compare the quality of the drawings in consideration and make decisions accordingly.

In an effort toward this need, we propose an overall quality measure of layout. This measure takes into account individual aesthetic criteria and gives a single numerical value. In this paper, we first briefly review current practices of quality evaluation of graph drawings. This is followed by an explanation of how our proposed measure is formulated and computed. Then, we present a user study for the validation of the new measure. This study has two sets of drawing stimuli. The first set of drawings are used to test its sensibility, while the other set are to demonstrate its capacity in predicting the performance of human graph comprehension. The paper concludes with a general discussion.

The main contributions of this work include:

1. A new overall quality measure was proposed based on the normalized *z* scores of individual aesthetics.
2. Empirical evidence was provided that demonstrates the sensibility and the predictive capacity of the new measure.
3. We found that the proposed measure was more sensitive to quality changes than traditional performance measures.

2. Related work

A typical graph visualization system includes two basic components: graph drawings and interaction methods that are intended to present these drawings in specific ways so that information embedded in the drawings can be processed effectively and efficiently by human users. There is a growing body of work on graph evaluation appearing in the literature [1,9,17,18]. This body of work can be divided into three major categories: system evaluation (including interfaces), interaction evaluation and graph drawing evaluation. In this section, we selectively review *evaluations of graph drawings with a focus on quality measurement*. Studies that contribute to a direct measurement of overall quality are also reviewed.

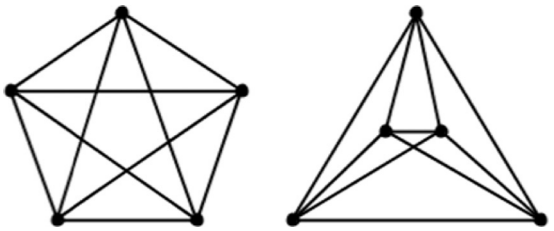

Fig. 1. An example of conflicting aesthetics. The left drawing shows maximum symmetries, but with more crossings, while the right one has minimum crossings, but with less dimensions of symmetry.