The Perception of Graph Properties in Graph Layouts

Utkarsh Soni[1], Yafeng Lu[1], Brett Hansen[1], Helen C. Purchase[2], Stephen Kobourov[3] and Ross Maciejewski[1]

[1]School of Computing, Informatics, and Decision Systems Engineering, Arizona State University, US [†]
[2]School of Computing Science, University of Glasgow, Scotland [‡]
[3]Department of Computer Science, University of Arizona, US [§]

Abstract

When looking at drawings of graphs, questions about graph density, community structures, local clustering and other graph properties may be of critical importance for analysis. While graph layout algorithms have focused on minimizing edge crossing, symmetry, and other such layout properties, there is not much known about how these algorithms relate to a user's ability to perceive graph properties for a given graph layout. In this study, we apply previously established methodologies for perceptual analysis to identify which graph drawing layout will help the user best perceive a particular graph property. We conduct a large scale (n = 588) crowdsourced experiment to investigate whether the perception of two graph properties (graph density and average local clustering coefficient) can be modeled using Weber's law. We study three graph layout algorithms from three representative classes (Force Directed - FD, Circular, and Multi-Dimensional Scaling - MDS), and the results of this experiment establish the precision of judgment for these graph layouts and properties. Our findings demonstrate that the perception of graph density can be modeled with Weber's law. Furthermore, the perception of the average clustering coefficient can be modeled as an inverse of Weber's law, and the MDS layout showed a significantly different precision of judgment than the FD layout.

1. Introduction

Given a particular graph, there are multiple graph drawing algorithms for producing a graph visualization. These algorithms remove edge crossings, depict symmetric substructures, and organize vertices and edges according to various heuristics and optimization techniques. Each graph drawing algorithm attempts to improve a user's ability to interpret a graph [Pur97] through its own optimization criteria and computational method. Experiments that compare the performance of different graph layout algorithms typically consider the visual properties of the graph drawings (e.g. vertex size [LLW16]), the extent to which common aesthetics are emphasized [Pur98], or their computational complexity. However, graphs can be characterized in many different quantitative ways, not only the number of vertices and edges, but also the structure of the graph (e.g., density, diameter, clustering coefficient, degree distribution etc.). To date, most comparisons between graph layout algorithms have focused on theoretical properties such as the computational complexity of an algorithm, but little work has explored the perception of graph properties with respect to a graph layout.

In this paper, we present the results of two experiments that focus on comparing graph layout algorithms with respect to the ex-

tent to which they support the perception of underlying properties of the graph. We hypothesize that some graph layout algorithms may make it easier for a viewer to discern graph properties than others. Following the example of perception studies in psychology and the work of Rensink et al. [RB10], we focus on measuring the just noticeable difference (JND) – that is, the smallest difference between two property values that can be perceived by humans – for two graph properties across three different graph layout algorithms. Rensink et al. [RB10] used a JND approach for the perception of correlation in scatterplots, showing that such perception can be modeled using Weber's law – that is, the JND between the perception of a given correlation (the target stimulus) and a different one is a constant ratio of the original target stimulus. Harrison et al. [HYFC14] used this JND methodology as a means of comparing nine different ways in which correlation can be visualized (e.g. scatterplot, parallel coordinates, donut charts). Subsequent work by Kay and Heer [KH16] revisited the data collected by Harrison et al. [HYFC14] and enhanced the analysis with log transformation and censored regression to inclusively embrace all individual data points. This same methodology offers a means for comparing graph layout algorithms with respect to perception of graph properties, providing a quantifiable means of determining which algorithms outperform others with respect to perception of important properties. For example, if a communication network is to be drawn so that people can easily discern the density of the network – which is the best layout algorithm to use?

[†] [1]{usoni1, lyafeng, brett.hansen, rmacieje}@asu.edu

[‡] [2]Helen.Purchase@glasgow.ac.uk

[§] [3]kobourov@cs.arizona.edu

We apply the experimental methodology of Rensink et al. [RB10] to determine if the perception of graph properties can also be modeled following Weber's law. As in Harrison et al. [HYFC14], we adopt this methodology for a crowdsourcing environment on Amazon's Mechanical Turk, and applied it to two graph properties (graph density and average local clustering coefficient) across three different graph layout algorithms (Force Directed - FD [Hu05], Circular [DMM96], and Multi-Dimensional Scaling - MDS [BP09]). In this way, we are able to:

- Demonstrate that the perception of graph properties can (sometimes) be modeled using Weber's law;
- Provide a means for comparing the effectiveness of different graph layouts for perceiving graph properties.

This work is important because graph drawings are increasingly being used in a variety of non-research areas (e.g., fraud detection, criminal networks, marketing) and depicting them in a way that makes the salient properties easy to perceive will make them more useful for the appropriate domain task. All stimuli and responses used in this experiment can be retrieved at http://vader.lab.asu.edu/GraphAnalytics/.

2. Related Work

Here we review related work in graph layout algorithms, graph mining and graph properties, and the perception of properties.

2.1. Graph Layout Algorithms

Research on the best way to automatically visualize relational data with node-link diagrams has been active for several decades. The early primer by Di Battista et al. [DETT99] provides a good overview of the field of graph layout algorithms, and researchers in this area continue to develop new approaches and improve on existing ones. More recent are comprehensive surveys of the area by von Landesberger et al. [VLKS*11] and Gibson et al. [GFV12].

Graph layout algorithms fall into different categories. Force-Directed algorithms [Kob13] are based on a physical springs model and attempt to achieve minimum energy. Dimension Reduction algorithms project high-dimensional data onto low dimensional space, typically preserving the graph-theoretic distance between pairs of vertices [CB09]. Orthogonal layout methods insist that all edges are horizontal or vertical [EFK01]. Circular algorithms place all vertices on the circumference of a circle in careful ordering so as to reduce edge crossings [DMM96]. Other graph layout techniques have also been proposed to facilitate fast calculation of large networks and human-centered designs. Galán and Mengshoel [GM18] proposed a neighborhood beautification layout technique which has each node passing messages to its neighbors to adjust their position. This approach gives a trade-off between the layout quality reached by force-directed methods and the fast runtime achieved by algebraic methods. HOLA [KDMW16] was proposed as an automatic network layout algorithm that learns aesthetic criteria from users and encodes these criteria to its layout. There are also specific algorithms for trees [RT81] and planar graphs [NR04]. Each graph layout algorithm attempts to optimize one or more specified criteria (sometimes called "aesthetics") in an attempt to make the graph

easy to interpret – for example, minimizing the number of edge crossings, or fixing vertices and edges to a grid [Pur97].

In this paper we use three layout algorithms commonly implemented across a variety of different graph drawing systems, and each algorithm is an example of a different category. The first one is a multi-level force-directed algorithm by Yifan Hu [Hu05], available as sfdp (scalable force-directed placement) in GraphViz [EGK*01] and as YifanHu in Gephi [BHJ*09]. We use the implementation provided in Gephi. The second algorithm is one of the most effective dimensionality reduction algorithms, as shown in an experimental study by Brandes and Pich [BP09]. This algorithm applies classical multi-dimensional scaling, followed by stress majorization. We use the implementation provided in the MDSJ library in Gephi (https://gephi.org/plugins/#/plugin/mdslayout). The third algorithm is a circular layout [DMM96]. It places all vertices evenly spaced along a circle and attempts to reduce the number of crossings. We use the implementation provided in Gephi. The first two algorithms chosen are representative of the major methods (force-directed and MDS respectively), while the third algorithm (circular) is a good "generic" option provided by most graph drawing systems. Other algorithms and layout categories should be explored in the future; however, in this work we focus on these algorithms due to their ubiquity in freely available graph drawing systems.

2.2. Graph Properties and Graph Mining

The structure of any graph can be characterized by a set of graph-theoretic properties, the most trivial being order (number of vertices) and size (number of edges). These properties describe the nature of a graph, independent of the way it is drawn – other examples include density (the proportion of the number of edges to the maximum possible), diameter (the longest shortest path between pairs of vertices), number of connected components (subgraphs with a path between any pair of vertices), vertex degree distribution (represented as a histogram), and clustering coefficient (the extent to which vertices are clustered together by edges).

Graph mining (as a type of structured data mining) is the activity of identifying patterns in graphs. Rehman et al. provide a comprehensive review of graph mining approaches [RKF12]. Chakrabarti and Faloutsos address the issue of generating synthetic graphs that match the patterns within real-world (especially large) graphs [CF06]. In particular, they emphasize the importance of being able to say that two different graphs are similar to each other with respect to given properties. The PEGASUS system handles very large graphs (with "billions" of vertices), to find connected components, diameter, and vertex proximities [KTF09].

Visualization is commonly used for depicting the values of graph properties, although this is not always using the common node-link diagram. Kairam et al. describe GraphPrism [KMSH12], which visualizes several graph properties using stacked histograms and color encoding, including connectivity, transitivity, and density. Their user study asked participants to choose a synthetic graph (from a set of ten, all represented in the GraphPrism diagram format) that best matched the summary statistics of a real graph. Visual analysis of large graphs represented by adjacency matrices is considered by van Ham et al. [VHSD09]. The tabular interactive tool

ManyNets [FPSG10] can visualize a large number ("up to several thousand") of graphs, allowing visual comparison based on properties including degree distribution and clustering coefficient. The Visone system [BW04] integrates both analysis and visualization of social networks, focusing on properties associated with the structural importance of vertices (e.g. in-degree, closeness, authority). Vizter [HB05] also produces node-link drawings of graphs, in this case identifying and highlighting communities in social networks. More recent and ongoing work also considers whether the aesthetic criteria of graph layouts contribute to human understanding of the relational data [HEH14, KPS14].

This prior work reinforces the need for properties of graphs to be made evident to users – either through a supplementary visualization (as in ManyNets [FPSG10]), or as part of the node-line depiction (as in Visone [BW04]). Being able to compare two graphs according to their properties is particularly important when synthetic graphs are to be used in place of larger real-world ones [CF06], but no studies so far have considered this question from an empirical human perception perspective. Specifically, can humans detect differences in the properties of graphs when depicted as graph drawings, and are there particular layout algorithms that best support the visual perception of graph properties?

2.3. Perception of Properties: Just Noticeable Difference

Experiments in the psychology of perception use the "Just Noticeable Difference" (JND) as a means of determining the minimum distinguishable property difference between two stimuli. Such experiments ask participants to indicate which of the two stimuli has a greater value of a given measurable property (e.g., which square is greener). The JND is the value difference between two stimuli that is noticed at least 50% of the time by participants. The JND experimental method can be applied to any type of perception. For example, Goodfellow [Goo34] studied sound, vision and touch, Mahy et al. [MEO94] considered color, and Wilson et al. [WHBH11] studied thermal perception. Weber's law states that if P is the property value of a stimulus, the ratio of the JND to P will be constant.

While Weber's law experiments typically focus on low-level perceptual properties, they can be applied to any stimuli for measurable properties. For example, de Silva et al. [DSFW*10] looked at the 3D stereoscopic vision, Cornelissen et al. [CGCT16] investigated visual biases in the perception of body weight, Camacho et al. [CDGS15] considered the perception of viscosity in beverages, and Chowdhury et al. [CCS15] looked at travel time and route taking decisions. Datasets also have properties; for example, a set of numbers has a mean, median, and standard deviation, and two variables have a correlation coefficient [RB10]. Harrison et al. [HYFC14] and Kay and Heer [KH16] used a JND approach for correlation coefficient, as shown in scatterplots and other visualizations.

We apply the same JND methodologies for the perception of graph properties. We aim to determine whether the perception of two graph properties (graph density and average local clustering coefficient) follow Weber's law, and which of the three graph layout algorithms (FD, MDS, and circular) best supports the perception of property difference.
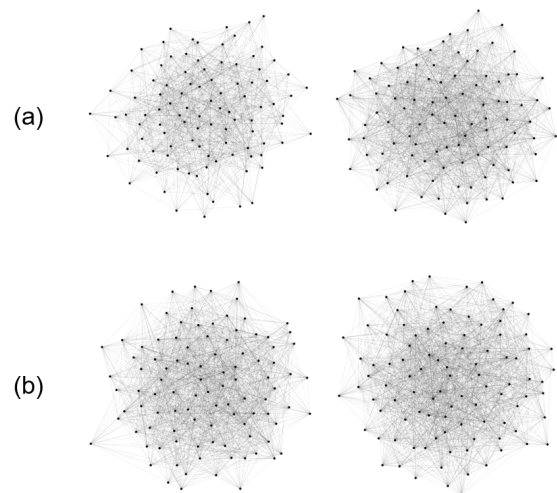
Figure 1: *(a) A sample starting comparison with target value $d = 0.2$ on the left and $d = 0.3$ on the right. Participants were asked to choose which one has a higher graph density. (b) The staircase procedure converges to the JND by gradually making comparisons more difficult: $d = 0.3$ on the left and $d = 0.28$ on the right.*

3. Methodology

The goal of this work is to quantitatively evaluate human perception of selected graph properties and compare the way in which different graph layout algorithms support the perception of these properties. We conducted five user studies to quantify the just noticeable differences (JNDs) of two graph properties and measured how these JNDs fit as a function of the property value in three different layout contexts. We hypothesize that some graph layout algorithms will be better than others at revealing graph properties.

In order to test this hypothesis, we chose two graph properties, graph density (GD) and average local clustering coefficient (ALCC), and three graph layout algorithms, Force Directed Layout [Hu05], Circular [DMM96], and Multi-Dimensional Scaling [BP09]. Graph density was chosen due to its simplicity to explain to participants and importance in expressing the connectedness of a graph. Previous studies have also explored the ability of humans to perceive density within plots [CM*85] indicating that (minimally) perception of graph density should be measurable. Average local clustering coefficient is a measure of the degree to which vertices in a graph tend to cluster together and is commonly analyzed with respect to small world networks. Given the recent importance of visualizing small world networks in the context of social network graphs, average local clustering coefficient was chosen as our second property to measure.

We performed two experiments. Experiment 1 consists of three user studies analyzing the perception of graph density over three layout algorithms. Experiment 2 consists of two user studies anayzing the perception of the average local clustering coefficient over two layout algorithms.

3.1. Experimental Method

Based upon the perceptual analysis of correlation done by Rensink and Baldridge [RB10] and Harrison et al. [HYFC14], we applied the same adaptive psychophysical method, a staircase procedure, to derive JNDs for the perception of the graph properties. For each user study, a list of evenly separated property values in a possible value range is designed as *base values*. Each base value graph drawing is compared to a graph drawing with another property value using two approaches (*above* or *below*). This means, for each property base value, the JND will be approached from above and below in two sets of comparisons using a staircase procedure.

To evaluate JND, participants are shown two stimuli side-by-side (in this case node-link graph drawings generated by the same algorithm), and participants will be asked to indicate which graph has a higher property value (see Figure 1 as an example of graph density). One of the two graph drawings has the base value, and the other graph drawing represents a smaller property value in the *below* approach or a larger property value in the *above* approach. Initially, the difference in the value of the properties between the two stimuli is set by a specified *initial-difference*. There is also a *step-size* parameter which is tuned to adjust the two stimuli in the staircase procedure. For example, if a participant is assigned the base value of 0.5, with a *below* approach, and the initial-difference is 0.1 and the step-size is 0.01, the first pair will have property values of 0.5 and 0.4. The participant is asked to select the graph drawing with the higher value of the property. Using a similar staircase methodology to Harrison et al. [HYFC14], if the correct choice is made (0.5), the next pair presented will be 0.5 and 0.41 (a decrease of 0.01 of the difference, making the task more difficult). If the participant answers incorrectly (0.4), the next pair will be 0.5 and 0.37 (an increase of 0.03 of the difference, making the task easier). A similar procedure is followed for the above approach. The distance changes allow the process to converge to a state such that the difference in properties between two side-by-side stimuli can be discriminated 75% of the time.

The staircase procedure ends when it reaches one of the following two conditions: (1) the participant has done a maximum number of judgments (e.g, 50), or (2) the participant reaches the JND indicated by a convergence criterion. This convergence criterion is the same as that of Harrison et al. [HYFC14] and Rensink and Baldridge [RB10]. Specifically, the convergence criteria uses the last 24 judgments to determine if the participant's ability to discriminate between the two property values from the given graph drawings has stabilized. To test the stability, these 24 judgments are segmented into 3 groups of 8 sequential judgments in each, and an F-test($F(2,21), \alpha = 0.1$) is applied on these 3 groups. When the F-test shows no significant difference between these three groups, convergence is assumed and the staircase procedure ends. No matter which ending condition a participant reaches, the final JND of the base value and the approach (*above* or *below*) is calculated by taking the average of the difference between the stimuli over the final 24 judgments.

For each trial in the experiment, the location of the base value graph is randomized (i.e., the base value graph will randomly be the left image or the right image). For each property value (both the base values and the interim values), we created a large number of possible graph drawings to choose from, thus mitigating against any possible learning effect or unanticipated confounding factors. All graph drawings are pre-computed and images used are chosen through random selection from our pre-generated graph drawing pool. The same methodology is applied to both the graph density and the average local clustering coefficient experiment. Experiments differ only in the choice of base values, initial-difference, and step-size which were designed after preliminary experiments were conducted to determine feasibility.

3.2. Data Analysis Method

Once data is collected, our goal is to determine if the data can be modeled using Weber's law. Prior to model fitting, we analyzed the data to remove any base value and approach conditions that suffer from the ceiling effect, which means the obtained JND is constrained by the range of the property value available in our experiment and therefore we could not observe the true JND. The ceiling effect is quantified in our experiment over the last 24 judgments. If over 50% of the judgments are performed on data within .05 of the upper or lower bounds of the data range, then participants are bounded by a ceiling effect. The *hit rate* is the percentage of participants that are bounded by a ceiling effect within a group (value, approach). Outliers outside of 3 median absolute deviations from the median in each base value and approach condition are also removed prior to model fitting.

After data cleaning, results from the experiment were analyzed in a three-step process. First, the JNDs in each user study (where an user study consists of a single graph layout algorithm being tested for a single graph property) are tested by fitting a model of Weber's law. The fitting methodology is modeled on that of Rensink and Baldridge [RB10] and Harrison et al. [HYFC14]. Average JNDs are calculated for each base value and approach condition. The property values used in the model are adjusted from the base value by adding (or removing) $0.5 \times JND$ within each base value and approach condition. To test for Weber's law, the JNDs and the adjusted values are fit using a linear regression model. Next, individual data points are fit without averaging and property value adjustment using linear regression with both continuous and categorical variables, following the methodology by Kay and Heer [KH16]. This is followed by an analysis to test if some data transformation is required for an adequate model. Finally, results are compared between layouts using Mann-Whitney U test to see if the perception ability in different layout algorithms has comparable distributions.
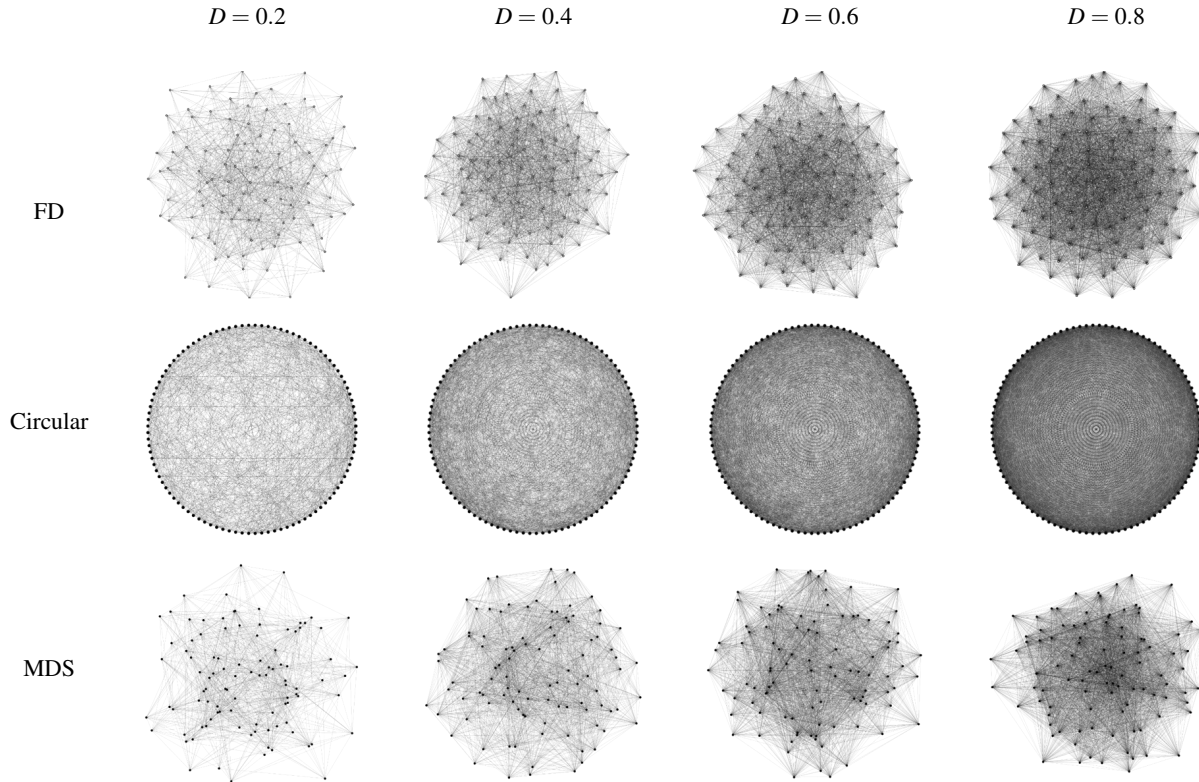
4. Experiment 1: graph density

Our first experiment explores the effects of graph layouts on perceiving graph density. In this experiment, we only consider undirected graphs where graph density is defined as:

$$D = \frac{2|E|}{|V|(|V| - 1)} \qquad (1)$$

where E is the number of edges and V is the number of vertices in the graph. Graph density describes how dense a graph is, and, for a fixed number of vertices, the more edges a graph has, the higher the graph density value. In a simple and connected graph,

Table 1: *Examples of the graph drawings by the three layout algorithms at several graph density values.*

	$D = 0.2$	$D = 0.4$	$D = 0.6$	$D = 0.8$

the maximum D is 1 (for complete graphs) and the minimum D is $\frac{2}{|V|}$ (for having $|V|$ - 1 edges connecting all vertices).

In our experiment, the goal is to quantitatively analyze people's perception of graph density given different layout algorithms. To achieve this goal, mixed design experiments were conducted such that just noticeable differences (JNDs) were collected by asking participants to compare the graph drawings of two different graph density values. We conducted three user studies with three different graph layout algorithms (FD, Circular, and MDS). Note that the formula of graph density we used here is nonlinear with respect to the graph order. Changing the number of nodes could result in a different perceptual model [Mel06]. In this experiment, we only study graphs with 100 nodes, and future work should explore how perceptual discrimination responds as the graph size changes.

4.1. Graph Generation

To study the perception of graph density, graphs with varying density need to be generated and visualized. Considering the size of the display and potential cluttering issues of node-link diagrams, we chose our graph order to be $|V| = 100$ in all experiments. Therefore, a simple connected graph has its graph density range from 0.02 to 1. Graphs are generated for every value in this range at intervals of 0.01, resulting in 99 different graph density values.

We generate a simple, connected graph G with graph density D stochastically by an iterative procedure. The initial graph has 100

vertices and 99 edges connecting the vertices in a path. At each step, a vertex is randomly selected and connected to another vertex that is not already its neighbor. This increases the edge count and thus the graph density. This process is repeated until the graph density comes within the range $D \pm T$. Here T is a tolerance parameter to define the accuracy of the output graph density and the value we use is 0.001. Algorithm 1 thoroughly describes this procedure.

Due to its stochastic nature, Algorithm 1 generates different graph structures for same graph density D. Different graph structures with the same graph density may vary on other graph properties, which cannot always be directly controlled for in our study. To mitigate the impact of perception on other graph properties associated with one particular graph structure, 50 graph structures for each D are generated and will be used randomly in the experiment.

Once graph structures are generated, the layout algorithm is applied to create graph drawings. The following three layout algorithms in different categories are used in our study based on their popularity and ease of use: FD [Hu05], Circular Layout [DMM96], and MDS [BP09]. These layout algorithms also have random factors when they position the vertices. To mitigate this impact, we randomly create 20 graph drawings using each layout algorithm for every graph structure. Therefore, for graph density, we have 99,000 ($99(values) \times 50(structures) \times 20(layouts)$) graph drawings by each layout algorithm. This is the pre-generation of our graph drawing pool used in our experiment, and Table 1 lists some

Algorithm 1: Generate graph with given graph density

Input : Number of Vertices(N), Required graph density(D), Tolerance(T)

Output: connected simple graph G with graph density D

Create a path graph G with N vertices $v_1, v_2, \ldots, v_n$ having edges $\{v_i, v_{i+1}\}$ where $i = 1, 2, \ldots, n-1$. Let set of vertices be $V = \{v_1, v_2, \ldots, v_n\}$;

Calculate the graph density of G as D_{new} ;

while $|D - D_{new}| > T$ **do**

    Randomly select a vertex v_i with uniform probability ;

    Find its current neighbors N_i ;

    Randomly select a vertex v_k with uniform probability from the set $V \setminus \{N_i \cup v_i\}$;

    Add an edge $\{v_i, v_k\}$ to graph G ;

    Recalculate graph density d_{new} ;

end

Table 2: *Parameters in the Model of Weber's Law*

Property	Layout	β_0	β_1	R^2	r	RMS
GD	FD	.0277	.0402	.904	.95	.0026
	Circular	.0235	.0582	.832	.91	.0052
	MDS	.0337	.0261	.438	.66	.0059
ALCC	FD	.5763	-.6478	.823	-.90	.033
	Circular	–	–	–	–	–
	MDS	.3619	-.3796	.911	-.95	.013

example graph drawings by these three layout algorithms for different graph density values.

We recognize that the order of the graph (the number of vertices) may also relate to a participant's ability to perceive JND. However, as the order of a graph increases, many algorithms converge towards a hairball layout. The goal of this experiment was to use a constant (modest) graph order to evaluate the perception of graph properties with respect to layout. Future studies will explore the range of graph orders in which graph properties are perceptible.

4.2. Procedure

The experimental procedure is a mixed 7×2 design in which there are 7 base values (0.2, 0.3, 0.4, 0.5, 0.6, 0.7, and 0.8) and two approach conditions (*above* and *below*). For each base value, D, the JND will be estimated from *above* and *below*.

Graph density may not be a widely known term and it is possible that our participants on Amazon Mechanical Turk (AMT) may not know about this concept. As such, we present an introduction page describing the meaning of graph density and present example graph drawings with low and high density values. To verify that participants understand the concept of graph density, they have to pass a screening session with 20 judgments. These judgments are designed to be easy (difference value > 0.2 between the two stimuli) so as to not exclude participants based on their perceptual ability. The goal of the screening is only to ensure that they understand the concept and know what the task is. For the first 5 judgments, participants will receive feedback on their choice. If they choose incorrectly, they have to explicitly make the right choice before they can move to the next judgment. For the final 15 judgments of the screening, there is no feedback, and the participant has to make at least 10 correct choices to continue to the real session.

For each layout algorithm, the conditions in this experiment include the seven base values and two approaches (*above* and *below*). Each participant is randomly assigned two base values and both *above* and *below* trials will be conducted for each base value. This results in four trials per participant and each trial consists of at

most 50 judgments. After each judgment, the screen will flash gray to notify participants that a new set of images to be judged have been rendered. In practice, the experiment takes approximately 10 minutes to complete. Following the completion of all four trials, a demographics questionnaire was given to participants. Finally, a short debriefing is provided. Payout rates were $.50 per participant.

4.3. Results

We recruited 105 participants for the user study of Circular layout, 105 participants for the user study of the MDS layout, and 102 participants for the user study of the FD layout. For each user study, this yields 30 data samples for every base value and approach condition and 420 (408 for FD) data samples in total. Prior to analysis, we removed outliers that are outside of 3 median absolute deviations from the median in each base value and approach condition. 37, 12 and 21 samples were removed for FD, MDS, and Circular respectively ($< 10\%$). No group level ceiling effects were observed.

Among all the participants that shared their gender and age information post study, the age and gender distribution in the Circular, MDS and FD layout experiments were 60 females and 42 males(age varying from 18-68), 64 females and 39 males (age varying from 18-76), and 62 females and 34 males (age varying from 19-71) respectively.

4.3.1. The Model of Weber's Law

Following classical work on perceptual laws [RB10, Ste57], we average individual JNDs over (value, approach) groups and calculate the adjusted graph density value off the base value before fitting a regression model. The adjusted density value, D_A, of each base density value, D, is calculated by shifting towards the approach direction by half of the average JND of the group.

$$D_A = D + 0.5 \times a_i \times JND \tag{2}$$

$$a_i = \begin{cases} 1 & \text{if approach is from } above \\ -1 & \text{if approach is from } below \end{cases} \tag{3}$$

The average JNDs are further fit by the adjusted graph density values through a linear regression.

$$JND = \beta_0 + \beta_1 D_A + \varepsilon \tag{4}$$

The same modeling process is applied to the data collected in the three user studies about graph density with different layout algorithms. The model coefficients (β_0 and β_1), R^2, the root-mean-square (RMS) error, and the correlation (r) between D_A and JNDs
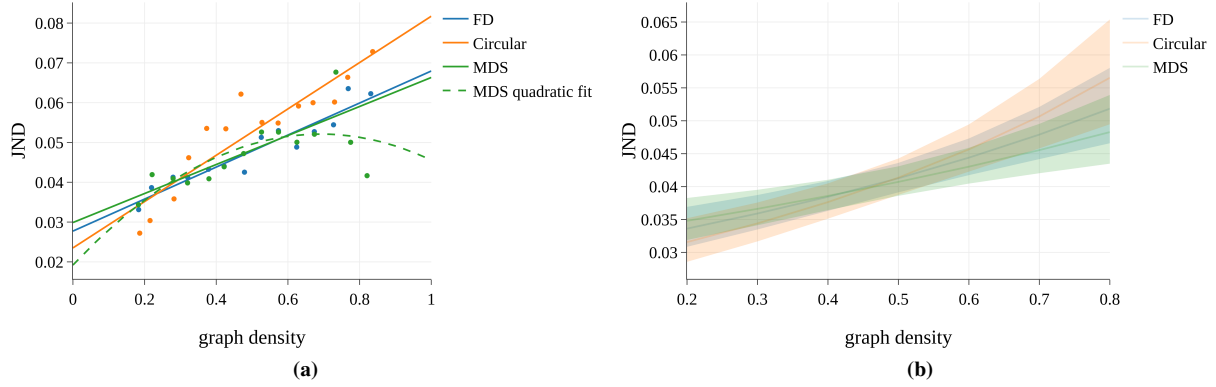
Figure 2: *Regression results for graph density user studies. (a) The model fit for the averaged individual JNDs. (b) The model fit for individual points after the Box-Cox transformation where the colored area indicates the 95% confidence interval.*

are listed in Table 2. Figure 2a shows the fit lines for the three layout algorithms along with the observed average JNDs. Among the three layout algorithms (FD, Circular, and MDS), both FD and Circular layout have a high goodness-of-fit ($R^2 \approx 90\%$ and $R^2 \approx 83\%$ respectively) and appear to follow Weber's law. For these layouts, participants were able to better discriminate between graphs when the density is lower and such ability decreases linearly when graph density increases.

While the perception of density in the FD and Circular layout follow Weber's law, we find that the linear model for the MDS layout only explains 44% of the variance. We see in Figure 2a (green) that when the graph density becomes large, the model fails to fit the underlying data. However, applying a quadratic regression to the MDS data results in a better goodness-of-fit ($R^2 = 0.53$), as shown in Figure 2a (green dashed line). Furthermore, calculating Cook's D [Coo77] for the MDS data, we find one leverage point of $D = 0.8$ and approach *above* (which has the largest Cook's D = 1.14, ~ 8 times the mean Cook's D of observations). After removing this observation, the goodness-of-fit of a linear model increased to $R^2 = 0.70$. This indicates that the MDS layout may follow Weber's law within a smaller range (specifically, [0.2, 0.7]) of graph density values.

4.3.2. Fitting Individuals

While we were able to fit models to the data, averaging individual JNDs could result in a loss of individual variance [KH16, CHM82]. As such, we have also analyzed the data following the approach of Kay and Heer [KH16], who re-analyzed Harrison et al.'s data to include individual variance. Taking the raw base value of graph density and the individual JNDs, we fit a linear regression model that includes the approach (*above/below*) as a categorical variable, a_i, which is defined in Equation 3. This model uses the raw base value of graph density, approach, and the interaction of these two

variables and is defined as:

$$JND_i = \beta_0 + \beta_1 \times D_i + \beta_2 \times a_i + \beta_3 \times a_i \times D_i + \varepsilon_i \quad (5)$$

To test the model's adequacy, we examined the residual distributions (Figure 3 (left)). By observation, we find that the residuals are skewed compared to a normal distribution, and a Box-Cox transformation for each dataset was applied using Equation 6. We then fit the model to the transformed data, Figure 2b.

$$JND_i^{(\lambda)} = \beta_0 + \beta_1 \times D_i + \beta_2 \times a_i + \beta_3 \times a_i \times D_i + \varepsilon_i \quad (6)$$

We take $\lambda = -0.5$ in our final model and this value is in the 95% confidence interval of the estimated λ for all three layout algorithms. Figure 3 (right) shows that residual distributions after this transformation become more normal. This indicates that the perception of graph density with the drawings given by the FD, Circular, and MDS layout algorithms may not follow an exactly linear relationship to the property value when individual variances are considered; instead, a power transformation may be required.

4.3.3. Comparison Between Layout Algorithms

In our experiment, workers are randomly recruited on AMT for each user study. In this way, subjects are considered as independent between conditions (layout algorithms) and the individual JNDs are independent measures. To compare the effect of the three layout algorithms (FD, Circular, and MDS) with respect to their ability to discriminate on graph density, the individual JNDs (as opposed to the mean JNDs) are used and following the work of Harrison et al. [HYFC14], a pairwise Mann-Whitney U test was applied. We applied Bonferroni correction for three pairs and set $\alpha = 0.0166$ in each test between two layouts. Results indicate that there are no significant differences in their individual JND distributions. Similarly, when the test is separately applied to the averaged JND, there was no significant difference found.

To further compare the three layout algorithms with respect to
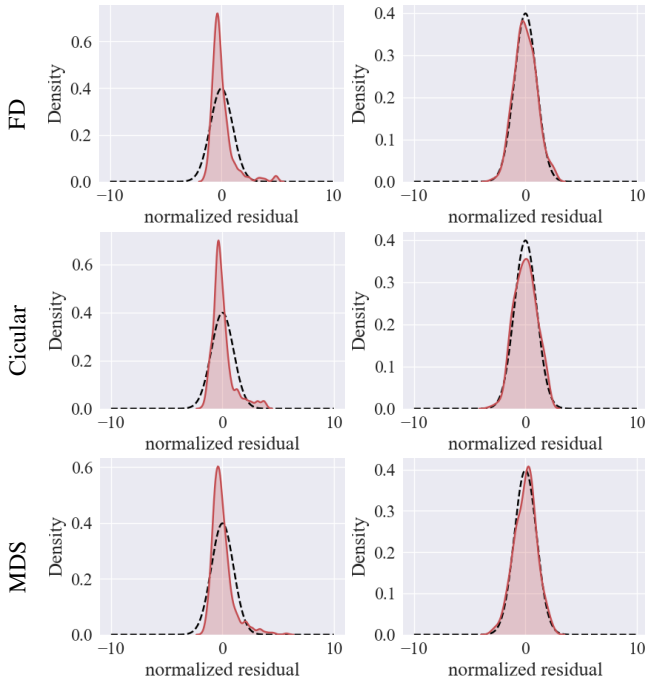
Figure 3: *The left shows that residual distributions were skewed before the transformation and the right shows the distribution is more Normal after the transformation.*

the discrimination of graph density, we examined the linear fit for the model of Weber's law (see Figure 2a). The clear overlap between all the layout algorithms for $JND < 0.55$ confirm our previous statistical findings that the three layout algorithm are not significantly different when perceiving graph density. Finally, we show our best model fit for all the individual points in Figure 2b where the colored area shows the 95% confidence interval of the model. This plot shows even more overlap across the entire range of D values confirming that all the three layouts are roughly equivalent for discriminating graph density.

5. Experiment 2: average local clustering coefficient

While graph density is a relatively straight forward property to visually explain, a primary aim for the second experiment was to begin exploring perception in graph layouts with respect to more complex graph properties. In real-world networks, vertices tend to create tightly connected groups and form clusters. This generates more clustering than random graphs. Along with scale-free property where degree distribution follows a power law, a high clustering coefficient is one critical characteristic of complex networks and plays an important role in graph analysis. As such, our second experiment focused on the perception of clustering in a simple, connected, undirected graph. In this experiment, we studied the perception of a global clustering measure, the average local clustering coefficient (ALCC), which is defined as:

$$C = \frac{1}{|V|} \sum_{i=1}^{|V|} \frac{2 \left| \left\{ e_{jk} : v_j, v_k \in V_i, e_{jk} \in E \right\} \right|}{k_i(k_i - 1)}. \qquad (7)$$

Algorithm 2: Generate Graph with tunable ALCC

Input : Number of Vertices (N), Number of edges for each vertex (m), Probability of Triad Formation (P_t)

Output: Connected simple scale-free graph with ALCC = C

Initialize graph G with small number of vertices, m_0, and no edges. Let V be the set of vertices of G, and k_v be the degree of vertex v ;

while *the number of vertices in $G < N$* **do**

    Add a vertex v with m edges to the graph G ;

    Select another vertex w of G with probability P_w which is proportional to its degree;

$$P_w = \frac{k_w}{\sum_{v \in V} k_v} \qquad (8)$$

    Add an edge connecting v and w ;

    while *unattached edges remain in v* **do**

        Perform the following TF step with probability P_t or PA step with probability $1 - P_t$;

        TF: add an edge between v and a neighbor of w;

        PA: select a new vertex w' other than v and w with probability, $P_{w'}$ given by Equation 8 and add an edge between v and w'. Update w with w'. ;

    end

end

This is the average of the local clustering coefficients of all the vertices measured by Watts and Strogatz [WS98] for 'small-world' analysis. In this equation, V is the vertex set and E is the edge set. V_i represents the immediately connected neighbors of a vertex v_i, and k_i is the degree of the vertex v_i.

In our experiment, we quantify how well people perceive clustering given graph drawings with different average local clustering coefficients and different layout algorithms. From an initial inspection among the three layout algorithms (Table 3), the circular layout with random vertex position provides little obvious discrimination between two graph drawings with large differences in ALCC. Therefore, the Circular layout is removed from this experiment. We conducted two user studies with the FD and MDS layout algorithms respectively for average local clustering coefficient.

5.1. Graph Generation

For this experiment, we generate graphs of $|V| = 100$ with varying ALCC while keeping the scale-free property. Specifically, we keep the number of edges the same across all graphs generated in order to keep graph density constant. However, by keeping the above properties, we can get only a limited range of ALCC value, C [HK02]. With $|V| = 100$ and $|E| = 194$, C ranges from 0.07 to 0.75. Note that the range of ALCC for any graph is 0 to 1.

We used the graph generation procedure of Holme et al. [HK02] that produces scale-free graphs with tunable clustering. It is an extension of Barábasi and Albert model (BA model) [BA99]. Algorithm 2 describes the procedure of our graph generation for varying C. In this algorithm, the Preferential Attachment (PA) step comes from the BA model and the Triad Formation (TF) step is an ex-

Table 3: *Examples of the graph drawings from the three layout algorithms at several average local clustering coefficient values.*

	$C = 0.1$	$C = 0.3$	$C = 0.5$	$C = 0.7$
FD				
Circular				
MDS				
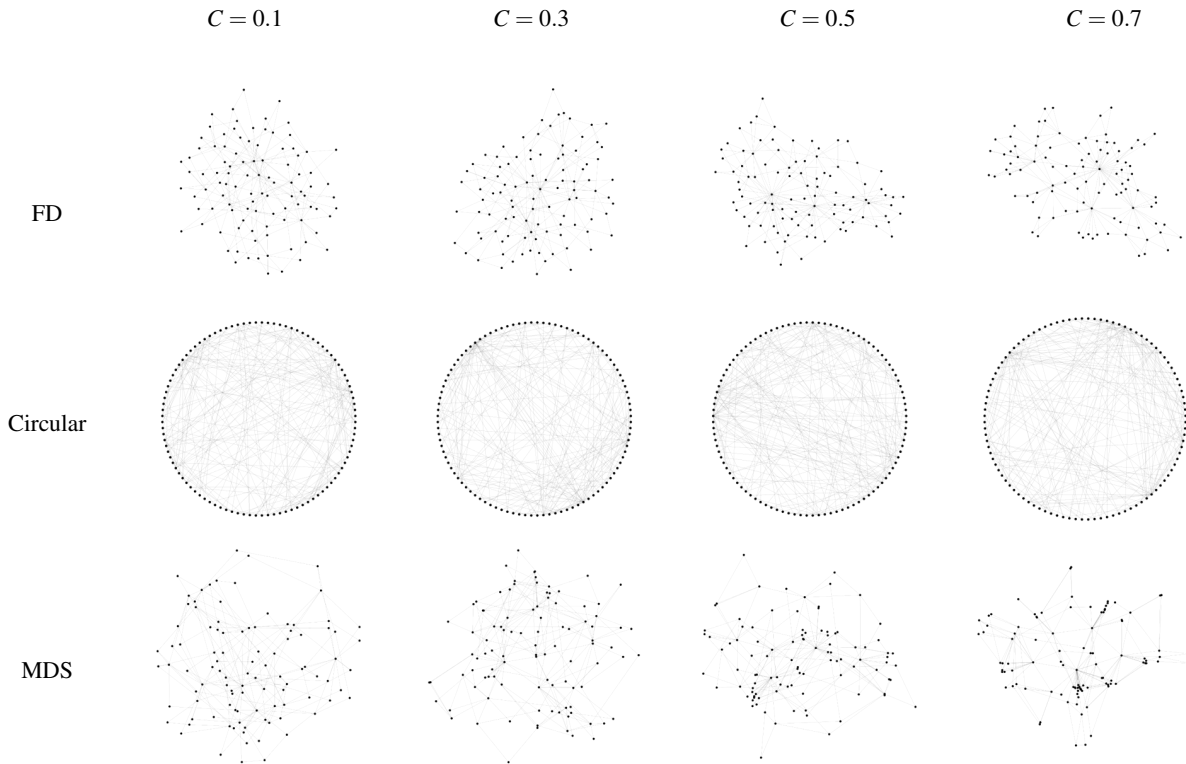
tension by Holme et al. [HK02]. Note that the PA step ensures the scale-free property of graph G, while the TF step increases C. Thus, by varying the probability P_t of taking a TF step, different values of C can be achieved.

We used Algorithm 2 with $m = 2$, and varied P_t to get graphs with different C. Similar to the graph generation for graph density, graphs are generated given every ALCC value, C, in the range of [0.07, 0.75] with step equals to 0.01. This results in 69 different average local clustering coefficient values. For each C, 50 graph structures are generated and each graph structure is visualized in 20 different graph drawings using the corresponding layout algorithm. Example graph drawings for varying C are shown in Table 3.

5.2. Procedure

Similar to Experiment 1, we present participants on AMT an introduction page to describe the meaning of clustering followed by two examples of graph drawings with high clustering and low clustering. Following the introduction, participants are required to take a screening session with 20 judgments which are designed to be highly discriminable. In each judgment, one graph has a low clustering value in [0.1, 0.3], and the other graph has a high clustering value in [0.5, 0.7]. For the first 5 judgments, the participants will receive feedback, and if the participant chooses incorrectly they have to explicitly choose the correct one to move on. For the final 15 judgments, there is no feedback and the participant has to make at least 10 correct choices to continue.

The experimental procedure for the average local clustering coefficient and MDS layout is identical to the procedure for graph density but with a different base value and step-size. This procedure has a 5×2 design in which there are 5 base values (0.2, 0.3, 0.4, 0.5, and 0.6) and two approach conditions (*above* and *below*) for each base value. The initial-difference is 0.1 and the step-size is 0.01, and the maximum number of judgments is set to be 50. For each average local clustering coefficient base value, C, the JND is estimated from *above* and *below* using the methodology presented in section 3.1. Each participant was randomly assigned two base values with both the *above* and the *below* approach.

The same procedure was applied for average local clustering coefficient and the FD layout. However, when collecting preliminary data for the FD layout, we found that participants suffered from a severe ceiling effect, which means the JND is bounded by the possible values we can generate. For example, we have $C = 0.07$ as the minimum average local clustering coefficient, and when we run (0.2, *below*), the furthest distance we can achieve is 0.13. This range is not discriminable by participants and prevents us from quantifying the the true JND for small values of C. Furthermore, during our preliminary data collection for the FD layout, we found that many of the participants completed all 50 judgments, which means that the experiment may have ended prior to the participant reaching a stabilized discrimination. Based on this information, we modified the base values and initial-difference for the full study. For the final FD experiment, we used a 10×2 design in which the base values for *above* and *below* are different. For the *above* ap-
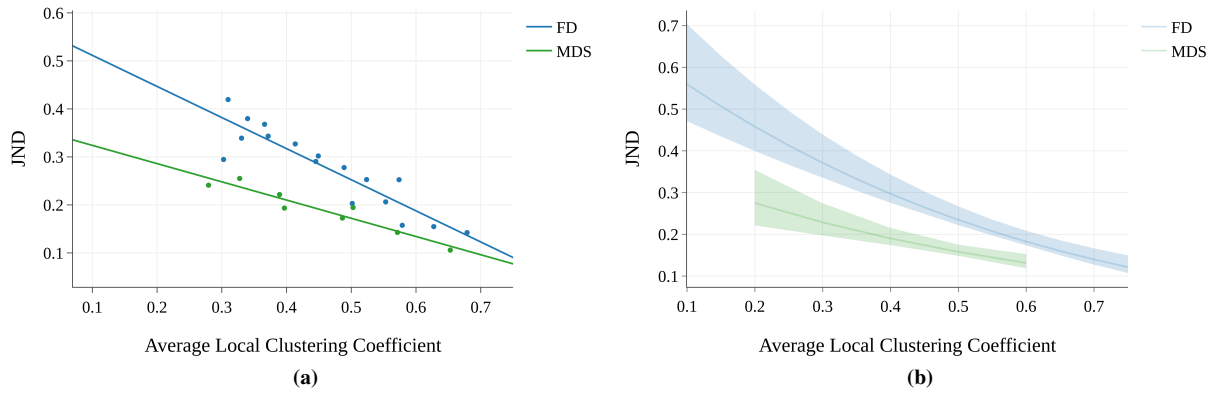
Figure 4: *Regression results for the average local clustering coefficient user studies. (a) The model fit for the averaged individual JNDs (b) The model fit for individual points after the Box-Cox transformation where the colored area indicates the 95% confidence interval.*

proach, lower base values (0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, and 0.55) are used, and for the *below* approach, higher base values (0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75) are used. To help participants reach their JND faster, the initial-difference is enlarged to 0.2. This value was chosen because the average JNDs of each base value we estimated in the Preliminary FD study either reached the ceiling or was larger than 0.2. Finally, we increased the maximum number of judgments to 75 to help participants reach a stable discrimination state.

5.3. Results

For the MDS layout, we recruited 75 participants. Each participant was assigned two base values with both *above* and *below* approaches. This gives us 30 data samples for every (value, approach) pair and 300 data samples in total. For the FD layout experiment, we recruited 201 participants. Each participant was assigned two base values, one from each approach. This gives us 20 data samples for every (value, approach) pair and 402 samples in total. Among all the participants that shared their gender and age information post study, the MDS layout study had 43 females and 32 males (age varying from 19-69) while the FD study had 110 females and 88 males (age varying from 18-70) respectively.

Before we model the JNDs for the ALCC, we analyzed the distribution of JNDs for each (value, approach) condition to identify outliers as well as groups that suffer from the ceiling effect (hit rate > 50%). In the collected JNDs for the MDS layouts, two groups, ($C = 20$, approach = *below*) and ($C = 30$, approach = *below*), had a hit rate greater than 50% (80.42% and 62.08% respectively). We removed these samples from our analysis as their true JND is not accurately measured. No outliers (points that fall outside 3 median deviations for each (value, approach) pair) were found.

Following the same procedure for the FD layouts, we removed the observations for $C \in \{30, 35, 40\}$ with approach = *below*. After removing groups due to the ceiling effect, outliers that fall out-

side 3 median deviation for each (value, approach) pair were also removed. In total, 60 samples were removed because of ceiling effects and 1 sample was removed as an outlier. 341 samples of the 402 samples collected are used in our analysis.

5.3.1. The Model of Weber's Law

After the observations were removed under the conditions mentioned, we averaged the JNDs in each group (base value × approach) and adjusted each base value C, following the same analysis procedure as section 4.3.1. The linear regression model with the adjusted base value C_A was fit for the averaged JNDs for both the FD layout and the MDS layout. The fit coefficients (β_0, β_1), R^2 and RMS error, and the correlation between C_A and JNDs are listed in Table 2. Results indicate that MDS has a better goodness-of-fit ($R^2 \approx 91.1\%$) than FD layout ($R^2 \approx 82.3\%$), while the average perceptions of ALCC displayed by both layout algorithms follow Weber's law with a negative linear relationship of the property value as illustrated in Figure 4a.

5.3.2. Fitting Individuals

As in Section 4.3.2, we also modeled all individual points without taking the average of the JNDs in each group. First, a linear model with base value C, approach, and their interaction term was fit for individual JNDs. Then, the residuals were analyzed against normal distribution. To correct the skewness of the residual distribution, a Box-Cox transformation was applied ($\lambda = 0.2518$ with Confidence Interval $(0.08, 0.42)$ for FD, and $\lambda = -0.058$ with Confidence Interval $(-0.279, 0.165)$ for MDS). This indicates that a log transformation for MDS and a power transformation with $\lambda \approx 0.24$ could fit the data better than a linear model when individual variance is considered. Figure 5 shows that residual distributions after this transformation become more normal.
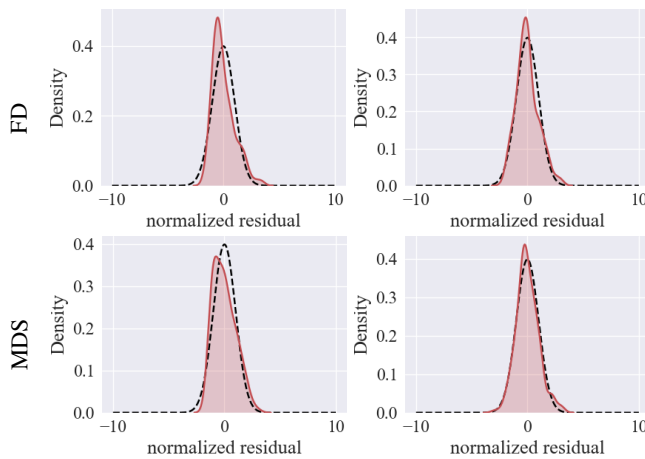
Figure 5: *The left shows that residual distributions were skewed before the transformation and the right shows the distribution is more Normal after the transformation.*

Table 4: *Common graph properties*

Property	Definition
Global Clustering Coefficient	It is given by $3 * n_\Lambda/n_\Lambda$ where n_Λ is the number of connected triplets (connected subgraph with 3 vertices and 2 edges) and n_Δ is the number of triangles.
Degree Distribution	Probability distribution of node degrees over the whole graph.
Average Path Length	Average of shortest distance between all pairs of vertices in a graph.
Assortativity Coefficient	Tendency of the vertices in a graph to be connected to other vertices with similar values of some vertex property (e.g., degree distribution).
Network Diameter	Greatest distance between any pair of vertices.

5.3.3. Comparison Between Layout Algorithms

Similar to section 4.3.3, we conducted a Mann-Whitney *U* test with $\alpha = 0.05$ to compare FD and MDS for presenting average local clustering coefficient. With p-value < 0.05 for both the test on individual JNDs and the test on the averaged JNDs, we find that there is a significant difference between the JNDs observed for the FD and MDS layout algorithm.

To further compare the two layout algorithms with respect to the discrimination of the average local clustering coefficient, we examined their best-fit linear regression models, Figure 4a. Here we see that the MDS layout algorithm provides a better perceptual discrimination of the average local clustering coefficient across the entire range of tested values.

Finally, we show our best model fit for all the individual points in Figure 4b with the colored area representing the 95% confidence interval of the model. This plot indicates that the MDS layout algorithm performs better then FD for perceiving the average local clustering coefficient.

6. Conclusion

To our knowledge, this is the first experiment designed to model humans' ability to perceptually discriminate graph properties. Such experiments provide us a means of quantitatively comparing graph layout algorithms with respect to their ability to communicate graph properties. The models and results presented in this work demonstrate that for the two graph properties tested, different layout algorithms can be modeled using Weber's law. Our experiment analyzed discriminations from 588 participants. Results in the perception of graph density showed that the three layout algorithms explored (Force Directed - FD, Circular, Multi-Dimensional Scaling - MDS) could be modeled using Weber's law and there was no significant difference between the layout algorithms. Results in the perception of average local clustering coefficient demonstrated that

the two algorithms considered (MDS and FD) can be modeled using Weber's law. This time there is a significant difference between the algorithms, as the MDS algorithm is better at discriminating ALCC than the FD algorithm.

In this study we explore only two graph properties and instances of graph layouts for three categories of algorithms. However, a large variety of graph properties remain to be studied; see Table 4 for a summary. Future work should explore more such properties and correlations between them, as well as more layout algorithms. An additional limitation is that all graphs tested in this experiment were of fixed order (number of vertices). As graphs become larger, some layout algorithms tend to produce hairball layouts. Future experiments should explore the effects of graph type, graph order, and screen size on perception to determine at what settings the discrimination of graph properties becomes infeasible, and how (or if) perception is affected. Furthermore, a comparison of perception within classes of layout algorithms should be done to explore if all layout algorithms that fall within a class (e.g., force directed algorithms) have the same basic underlying perceptual properties. Finally, more graph properties need to be explored. This experiment only scratches the surface of potential combinations. However, this serves as an initial step in demonstrating that (at least for some) graph properties can be discriminated. By identifying the different conditions and classes of algorithms that improve discrimination, future work can inform ideas of new design spaces for graph layout algorithms that not only focus on layouts for graph aesthetics, but also on conditions for graph perception.

7. Acknowledgments

References

[BA99]   BARABÁSI A.-L., ALBERT R.: Emergence of scaling in random networks. *Science 286*, 5439 (1999), 509–512. 8

[BHJ*09]   BASTIAN M., HEYMANN S., JACOMY M., ET AL.: Gephi: an open source software for exploring and manipulating networks. *International AAAI Conference on Weblogs and Social Media 8* (2009), 361–362. 2

[BP09]   BRANDES U., PICH C.: An experimental study on distance-based graph drawing. In *Graph Drawing* (2009), Springer, pp. 218–229. 2, 3, 5

[BW04]   BRANDES U., WAGNER D.: Analysis and visualization of social networks. *Graph Drawing Software* (2004), 321–340. 3

[CB09]   CHEN L., BUJA A.: Local multidimensional scaling for nonlinear dimension reduction, graph drawing, and proximity analysis. *Journal of the American Statistical Association 104*, 485 (2009), 209–219. 2

[CCS15]   CHOWDHURY S., CEDER A. A., SCHWALGER B.: The effects of travel time and cost savings on commuters' decision to travel on public transport routes involving transfers. *Journal of Transport Geography 43* (2015), 151–159. 3

[CDGS15] CAMACHO S., DOP M., GRAAF C., STIEGER M.: Just noticeable differences and weber fraction of oral thickness perception of model beverages. *Journal of Food Science 80*, 7 (2015). 3

[CF06]   CHAKRABARTI D., FALOUTSOS C.: Graph mining: Laws, generators, and algorithms. *ACM Computing Surveys 38*, 1 (2006), 2. 2, 3

[CGCT16] CORNELISSEN K. K., GLEDHILL L. J., CORNELISSEN P. L., TOVÉE M. J.: Visual biases in judging body weight. *British Journal of Health Psychology 21*, 3 (2016), 555–569. 3

[CHM82]   CLEVELAND W. S., HARRIS C. S., MCGILL R.: Judgments of circle sizes on statistical maps. *Journal of the American Statistical Association 77*, 379 (1982), 541–547. 7

[CM*85]   CLEVELAND W. S., MCGILL R., ET AL.: Graphical perception and graphical methods for analyzing scientific data. *Science 229*, 4716 (1985), 828–833. 3

[Coo77]   COOK R. D.: Detection of influential observation in linear regression. *Technometrics 19*, 1 (1977), 15–18. 7

[DETT99] DI BATTISTA G., EADES P., TAMASSIA R., TOLLIS I. G.: *Graph Drawing: Algorithms for the Visualization of Graphs.* Prentice Hall, Englewood Cliffs, NJ, 1999. 2

[DMM96]   DOĞRUSÖZ U., MADDEN B., MADDEN P.: Circular layout in the graph layout toolkit. In *International Symposium on Graph Drawing* (1996), Springer, pp. 92–100. 2, 3, 5

[DSFW*10] DE SILVA D. V. S., FERNANDO W. A. C., WORRALL S. T., YASAKETHU S., KONDOZ A. M.: Just noticeable difference in depth model for stereoscopic 3d displays. In *IEEE International Conference on Multimedia and Expo* (2010), IEEE, pp. 1219–1224. 3

[EFK01]   EIGLSPERGER M., FEKETE S. P., KLAU G. W.: Orthogonal graph drawing. In *Drawing Graphs*. Springer, 2001, pp. 121–171. 2

[EGK*01] ELLSON J., GANSNER E. R., KOUTSOFIOS E., NORTH S. C., WOODHULL G.: Graphviz - open source graph drawing tools. In *International Symposium on Graph Drawing* (2001), pp. 483–484. 2

[FPSG10] FREIRE M., PLAISANT C., SHNEIDERMAN B., GOLBECK J.: Manynets: an interface for multiple network analysis and visualization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2010), ACM, pp. 213–222. 3

[GFV12]   GIBSON H., FAITH J., VICKERS P.: A survey of two-dimensional graph layout techniques for information visualisation. 324–357. 2

[GM18]   GALÁN S. F., MENGSHOEL O. J.: Neighborhood beautification: Graph layout through message passing. *Journal of Visual Languages & Computing 44* (2018), 72–88. 2

[Goo34]   GOODFELLOW L. D.: An empirical comparison of audition, vision, and touch in the discrimination of short intervals of time. *The American Journal of Psychology 46*, 2 (1934), 243–258. 3

[HB05]   HEER J., BOYD D.: Vizster: Visualizing online social networks. In *IEEE Symposium on Information Visualization* (2005), IEEE, pp. 32–39. 3

[HEH14]   HUANG W., EADES P., HONG S.-H.: Larger crossing angles make graphs easier to read. *Journal of Visual Languages & Computing 25*, 4 (2014), 452–465. 3

[HK02]   HOLME P., KIM B. J.: Growing scale-free networks with tunable clustering. *Physical Review E 65*, 2 (2002), 026107. 8, 9

[Hu05]   HU Y.: Efficient, high-quality force-directed graph drawing. *Mathematica Journal 10*, 1 (2005), 37–71. 2, 3, 5

[HYFC14] HARRISON L., YANG F., FRANCONERI S., CHANG R.: Ranking visualizations of correlation using weber's law. *IEEE Transactions on Visualization and Computer Graphics 20*, 12 (2014), 1943–1952. 1, 2, 3, 4, 7

[KDMW16] KIEFFER S., DWYER T., MARRIOTT K., WYBROW M.: Hola: Human-like orthogonal network layout. *IEEE transactions on visualization and computer graphics 22*, 1 (2016), 349–358. 2

[KH16]   KAY M., HEER J.: Beyond weber's law: A second look at ranking visualizations of correlation. *IEEE Transactions on Visualization and Computer Graphics 22*, 1 (2016), 469–478. 1, 3, 4, 7

[KMSH12] KAIRAM S., MACLEAN D., SAVVA M., HEER J.: Graphprism: compact visualization of network structure. In *Proceedings of the International Working Conference on Advanced Visual Interfaces* (2012), ACM, pp. 498–505. 2

[Kob13]   KOBOUROV S. G.: Force-directed drawing algorithms. In *Handbook of Graph Drawing and Visualization*, Tamassia R., (Ed.). CRC Press, 2013, pp. 383–408. 2

[KPS14]   KOBOUROV S., PUPYREV S., SAKET B.: Are crossings important for drawing large graphs. In *International Symposium on Graph Drawing* (2014), pp. 234–245. 3

[KTF09]   KANG U., TSOURAKAKIS C. E., FALOUTSOS C.: Pegasus: A peta-scale graph mining system implementation and observations. In *IEEE International Conference on Data Mining* (2009), IEEE, pp. 229–238. 2

[LLW16]   LI J., LIU Y., WANG C.: Evaluation of graph layout methods based on visual perception. In *Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing* (2016), ACM, p. 90. 1

[Mel06]   MELANCON G.: Just how dense are dense graphs in the real world?: a methodological note. In *Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization* (2006), ACM, pp. 1–7. 5

[MEO94]   MAHY M., EYCKEN L., OOSTERLINCK A.: Evaluation of uniform color spaces developed after the adoption of CIELAB and CIELUV. *Color Research & Application 19*, 2 (1994), 105–121. 3

[NR04]   NISHIZEKI T., RAHMAN M. S.: *Planar graph drawing*, vol. 12. World Scientific Publishing Co Inc, 2004. 2

[Pur97]   PURCHASE H.: Which aesthetic has the greatest effect on human understanding? In *International Symposium on Graph Drawing* (1997), Springer, pp. 248–261. 1, 2

[Pur98]   PURCHASE H. C.: Performance of layout algorithms: Comprehension, not computation. *Journal of Visual Languages & Computing 9*, 6 (1998), 647–657. 1

[RB10]   RENSINK R. A., BALDRIDGE G.: The perception of correlation in scatterplots. In *Computer Graphics Forum* (2010), vol. 29, Wiley Online Library, pp. 1203–1210. 1, 2, 3, 4, 6

[RKF12]   REHMAN S. U., KHAN A. U., FONG S.: Graph mining: A survey of graph mining techniques. In *Seventh International Conference on Digital Information Management* (2012), IEEE, pp. 88–92. 2

[RT81]   REINGOLD E. M., TILFORD J. S.: Tidier drawings of trees. *IEEE Transactions on Software Engineering*, 2 (1981), 223–228. 2

[Ste57]   STEVENS S. S.: On the psychophysical law. *Psychological Review 64*, 3 (1957), 153. 6

[VHSD09]   VAN HAM F., SCHULZ H.-J., DIMICCO J. M.: Honeycomb: Visual analysis of large scale social networks. In *IFIP Conference on Human-Computer Interaction* (2009), Springer, pp. 429–442. 2

[VLKS*11]   VON LANDESBERGER T., KUIJPER A., SCHRECK T., KOHLHAMMER J., VAN WIJK J. J., FEKETE J.-D., FELLNER D. W.: Visual analysis of large graphs: state-of-the-art and future research challenges. 1719–1749. 2

[WHBH11]   WILSON G., HALVEY M., BREWSTER S. A., HUGHES S. A.: Some like it hot: thermal feedback for mobile devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011), ACM, pp. 2555–2564. 3

[WS98]   WATTS D. J., STROGATZ S. H.: Collective dynamics of 'small-world' networks. *Nature 393*, 6684 (1998), 440–442. 8