Social Relationship Visualization: The Best Choice Among Different Types of Charts

Student Name(s) & Student Number(s)：

Xiaosheng Liang (15211913)

Song Liu (15211324)

# Proposed Hypothesis (What Question Are You Asking?)

In this report, a user study on perception from data visualization is described, in which we measured the ability of participants to validate statements about the personal social relationship of data samples visualised using different chart types.

In assumption, hierarchy visualization (In this study, we use radial tree graphs) should be the most appropriate visualization type to illustrate the social relationship state of values for a variable.

In the field, several previous researches inspire us to construct the major idea of this study, such as “Friend Circles” [1] which is first introduced by Google+ and let users assign classmates, family members, colleagues and others to different group. Furthermore, the some recommendation of online social network are based on the “Trust Circles” [2] which is used to infer the friends from rating data concerning items that can be divided into different categories.

Take “Trust Circle” (Figure 1) as a example, the centre blue node represents ourselves, the green nodes represent the persons with whom you follow each other. The yellow nodes represent the people you followed but they don’t follow you. The red nodes represent the people who just follow you.

Additionally, we could analysis the different social network relationship from the overall view of trust circles in Figure 2.

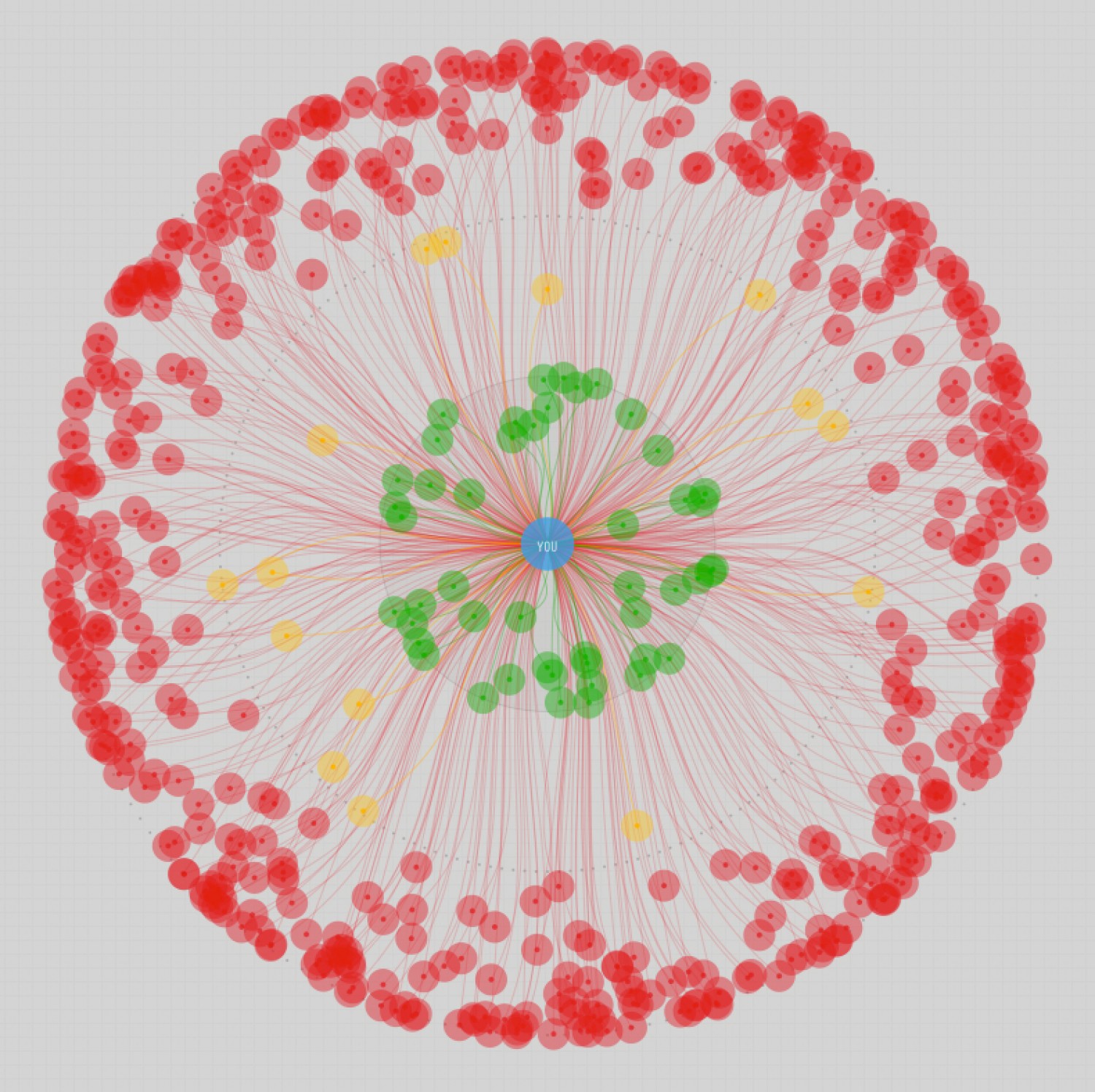
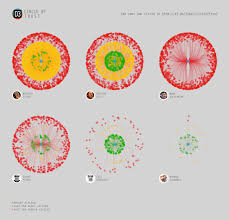
 

Figure 1 Figure 2

During experiments, the specific questions we would ask about this hypothesis are as follows:

Questions:

1. According to the information shown by the graphs, in which intra-person contacting frequency group the proportion of friends are the largest and which are the smallest?
2. According to the information shown by the graphs, who has the most frequent-contacting (over 30 times a month) friends than the others, and who has the most rare-contacting (less than once a month) friends?
3. According to the information shown by the graphs, who has more friends in touch and who has less recently (probably in hale a month)?
4. According to the information shown by the graphs, who is social type (have lots of friends who could be kept in touch with friends regularly), and who is unsocial?

These questions are essential due to the fact that they are well-designed for the purpose of detecting the difference between variant visualization functions on showing personal social relationship. These questions could be separated into two major parts: general or specific, in other words, statistic or detailed. In expected answer, only radial tree graphs could provide general information as well as detailed information effectively and efficiently. As for bar charts and pie charts, they could only show some kinds of general statistic information.

# Experimental Method

## Overview

**Dependent variables:**

Participants’ ability to read the details of social relationship visualization is measured by task complement time and accuracy.

**Independent variables:**

* Visual encodings used:

1. Radial tree graphs (3-D)

Dimension 1, *Distance from the centre*: contact frequency of friends (the further distance is, the less frequency they contacts);

Dimension 2, *Colour*: identify which frequency group does this friend belong to;

Dimension 3, *Area*: which frequency group occupy more area in this graph and how big the overall graph is.

2. Bar charts (2-D):

Dimension 1, *length of bars*: number of contacting friends for each frequency group;

Dimension 2, *x-axis and colour:* different social types and contact frequency

3. Pie charts(2-D):

Dimension 1, *angle (area)*: number of contacting friends for each frequency group;

Dimension 2, *Colour:* name of frequency groups.

* Colours used:

In radial tree graphs and pie charts, different colours are used to indicate different frequency groups; in bar charts, different colours are used to tell different social relationship types.

**Confounding variables:**

Prior expertise, learning effects and previous experiences.

To take part in the study, participants do not need to have any prior expertise in data analysis as we were interested in measuring the ability of average, non-expert viewers to interpret different chart types. We did restrict participation to US-based participants to control for English language capability. Additionally, different types of charts were shown to different people to eliminate learning effect.

**Experimental conditions:**

The light, noise and the other unrelated variables are under control during the experiment.

The combinations of chart types (3) and data from contact record of different social types (3) amounted to a total of 9 different conditions.

**Experimental design:**

Our experiments would be conducted through between-group design which means each group of subjects are assigned to one experimental condition. Tasks would be presented randomly to the participants in every group in order to remove learning effects. Furthermore, we learnt the experimental method from previous study in the field which study choosing the right charts to show data distributions to Non-Expert users [3], and designed our own experimental methods.

## Data collection

Participants in this experiment will be shown questions with sample paper at the same time. Their task completing time will be collected first. And then the accuracy of their answer will be manually measured. That is because the data collected from objective measurements is more specific and detailed to compare with and much more easy to figure out its social contacting type.

There is no mix between subjective and objective measurement, only objective measurements count in this experiment.

The most significant differences in these readability relates to the chart types. This set of well-designed tasks shows that a bar chart or a pie chart itself provides no ability to illustrate the social-related information well. To be specific, general statistic result with absolute value could be got easily by bar charts’ viewers and that with relative value could be got easily by pie charts’ viewers, however, the none of them are able to convey detailed social relationship like ’who and who is close friend’. With radial tree graphs, all of these issues are shown clearly.

## Selected subjects

In this experiments, 90 subjects would be used which include 45 male and 45 female. Considering between-groups method used, the experiment is required to use more subjects to ensure high accuracy. And they are divided into 3 groups which means 15 boys and 15 girls are involved in each group.

The criteria is that the subjects should have enough knowledge and experience to view the charts and interpret illustrated information, with about the same level of education. And they are not supposed to have any prior expertise or even perspective in social relationship analysis.

Considering the condition illustrated above, students from secondary school (age 12-16) who have enough knowledge and experience to view the charts and interpret illustrated information and without any prior expertise in data analysis would be the best choice.

The best way to source these subjects is to hold a public lecture among and secondary school students in the subject of social media. It would be a pretty attractive topic for them, and this experiment should be done before the lecture.

## Data analysis

Since three groups with same number of participants are employed by this experiment, the statistic test ANOVA should be used to analyse the result.

Before that, the average time of each group should be calculated. And the collected answers are going to be compared with the previously standardized answers, then calculate the accuracy of these answers. Given that 15 boys and 15 girls which are about the same age and same level of education (secondary school), taking the average performance is to make the subjects more representative.

Since then, the independent variables are tightly connected with the visualization function, so the analysis result of participants’ answer is directly related to the hypothesis ‘which is the best visual method to illustrate personal social relationship’.

## Practical setup

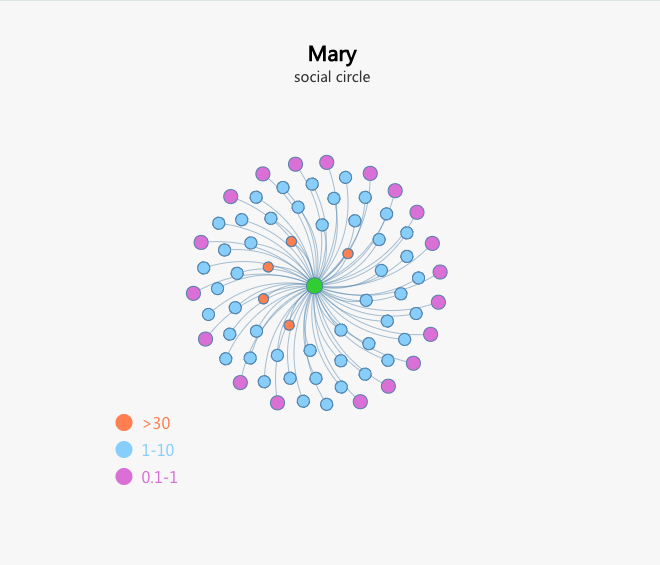
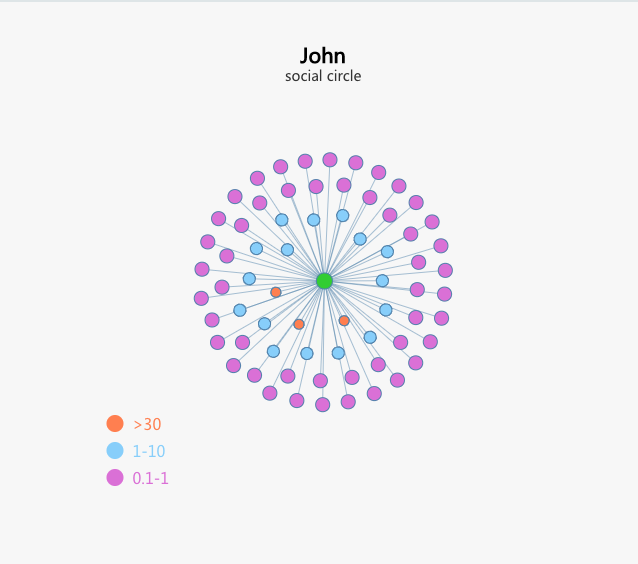
In this experiment, we would use offline method and the visualizations will be displayed on printed paper. This could be handled easily and it ensures every visualization samples are exactly the same.

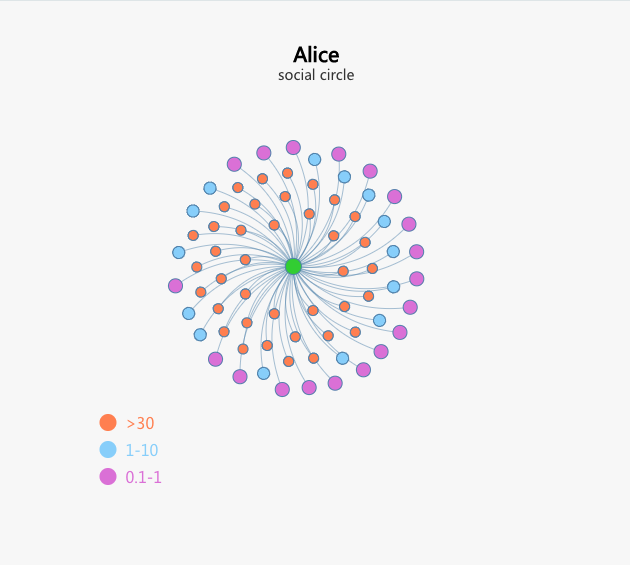
Furthermore, all the participants are supposed to take the test in the same room which provides the same environment like lights and noise.

# Data Visualisations

**The test for group 1:**

The central green node indicates the centre person who indicates different type of personal social relationship, such as Mary, John and Alice. The other nodes surrounding by that node represent the friends who contact to them. As for “0.1-1”, ”1-10”, “>30” presenting by different colours mean contact times to the person in a month.



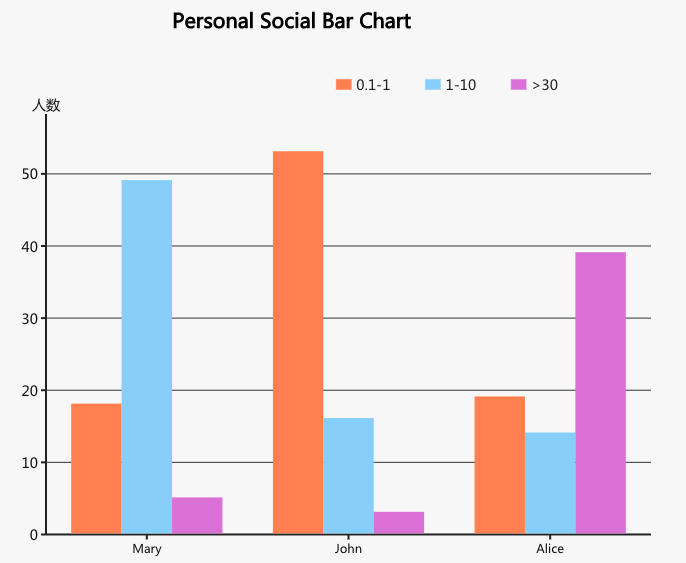
Please look the these three pictures and answer the question blew:

1. According to the information shown by the graphs, in which intra-person contacting frequency group the proportion of friends are the largest and which are the smallest?

1. According to the information shown by the graphs, who has the most frequent-contacting (over 30 times a month) friends than the others, and who has the most rare-contacting (less than once a month) friends?
2. According to the information shown by the graphs, who has more friends in touch and who has less recently (probably in hale a month)?
3. According to the information shown by the graphs, who is social type (have lots of friends who could be kept in touch with friends regularly), and who is unsocial?

**The task for group 2:**

Three different colours indicate contact times to the person in a month. The length of each bar describes the numbers of friends for each frequency group. As for Mary, John and Alice on x-axis illustrate different type of personal social relationship.



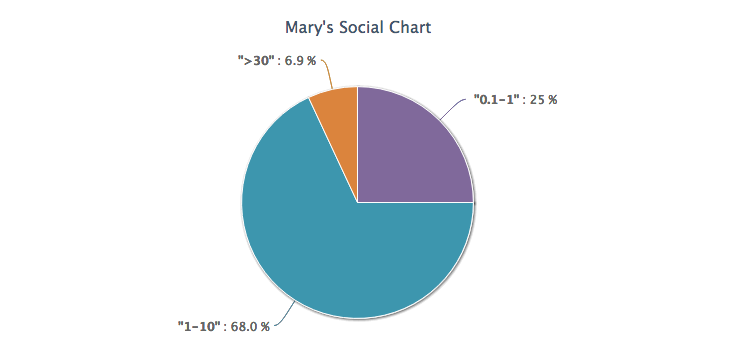
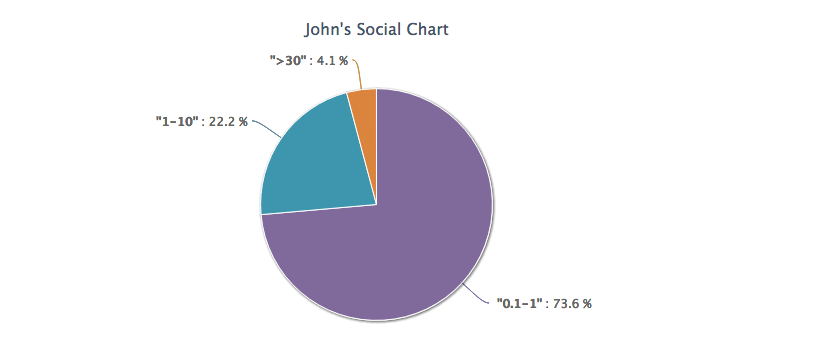
Please look the these three pictures and answer the question blew:

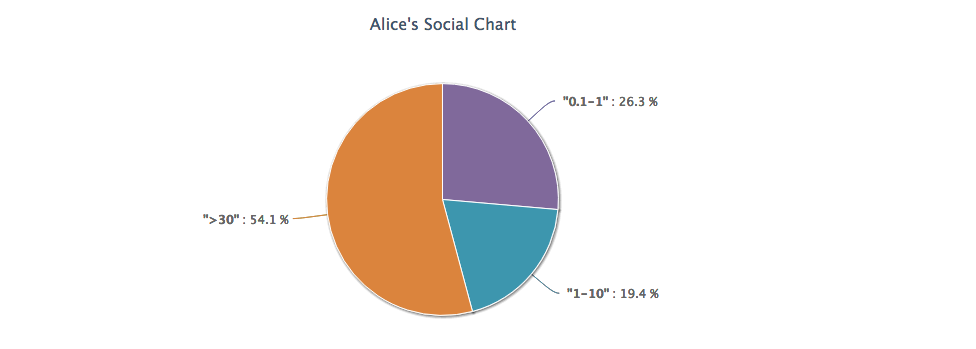
1. According to the information shown by the graphs, in which intra-person contacting frequency group the proportion of friends are the largest and which are the smallest?

1. According to the information shown by the graphs, who has the most frequent-contacting (over 30 times a month) friends than the others, and who has the most rare-contacting (less than once a month) friends?
2. According to the information shown by the graphs, who has more friends in touch and who has less recently (probably in hale a month)?
3. According to the information shown by the graphs, who is social type (have lots of friends who could be kept in touch with friends regularly), and who is unsocial?

**The task for group 3:**

Three charts describe different type of personal social relationship, such as Mary, John and Alice. The angles (areas) of each pie chart indicate the ratio of friends for each frequency group of all the friends. As for “0.1a-1”, ”1-10”, “>30” presenting by different colours mean contact times to the person in a month.



Please look the these three pictures and answer the question blew:

1. According to the information shown by the graphs, in which intra-person contacting frequency group the proportion of friends are the largest and which are the smallest?

2. According to the information shown by the graphs, who has the most frequent-contacting (over 30 times a month) friends than the others, and who has the most rare-contacting (less than once a month) friends?

1. According to the information shown by the graphs, who has more friends in touch and who has less recently (probably in hale a month)?
2. According to the information shown by the graphs, who is social type (have lots of friends who could be kept in touch with friends regularly), and who is unsocial?

# References

[1] Xiwang Yang, Harald Steck and Yong Liu. Circle-based Recommendation in Online Social Networks, 2012.

[2] Farhaan Mirza and Fernando Beltrán. Using an Agent-Based Friend Circle Creator Model to Analyze Drivers of Consumer Choice: Network Effects Vs. Value Proposition, 2013.

[3] Arkaitz Zubiaga and Brian Mac Namee. Knowing What You Don’t Know: Choosing the Right Chart to Show Data Distributions to Non-Expert Users, 2015.