

Final Project Proposal

Xander Fang, Jacob Lin

For the final project of this course, we want to implement the bubble set contour that is introduced in Collins, Penn and Carpendale's paper. The paper introduces techniques and algorithms to add contour layers above existing visualizations, such as scatterplots and tree diagrams. Set relation contours are a two-dimensional channel that serves to group different categories of data entries and provide spatial information for users. Instead of marking each single data points like other channels, such as color or shape, bubble set contours group data points at a higher level and view the group of data points as a whole, which makes them easier to catch attention and follow through.

We will implement the set relation bubble over scatterplots. In this paper, there are four versions of visualizations with set bubble implementations, but their visual effects differ from each other. We choose to implement set bubbles over scatter plots because we can implement multiple interactions along with set bubbles over scatterplots, which will have a better effect than other stationary versions.

We will use the dataset of Olympic athletes and results history from Kaggle (https://www.kaggle.com/heesoo37/120-years-of-olympic-history-athletes-and-results#athlete_events.csv) to implement our system. We will use the number of athletes who won medals as the Y-axis and the total number of athletes that participate in an Olympic game as the X-axis. In addition, the sizes of the dots in the scatterplot will be proportional to the male-female athlete ratio of each country. Besides, the colors of the dots represent which continent the country belongs to. This is just a temporary plan and is subject to change when we implement the system and learn more from the dataset.

Users will be able to interact with our system in multiple ways. Detailed information of each country will be animated when users click the corresponding dot. In addition, hovering on a country will make all non-members of the set be transparent which clarify the hovering set members. By doing that, it will be easier for us to discover the outlier members of the hovering set members.

As a team, we will split the work evenly and contribute as much as possible to the whole project. Xander will focus on pre-processing the dataset, and compute and draw the continuous bounding contours. Jacob will focus more on possibly build the basic scatter plot as well as designing the overall layout and CSS styling. As for the milestone, we will finish the basic layout and build the scatterplot, but we expect the set relation bubbles are going to take more time and we will complete it by the end of the quarter.