EmojiCloud - Write Up

Cheng Guo

I joined the EmojiCloud project when Dr. Feng invited me on February 7, 2022, to talk about creating a word cloud tool for emojis. I was generating the word cloud based on Python's word cloud package, which transfers all the emojis into the corresponding Unicode to plot. It cannot plot the correct color of the emojis and some emojis are rotated upside down. From several examples, including 🙂 🙃 and ❤️ 🧡 💛 💚 💙 💜 🖤 🤍 🤎, we learned that color and rotation matter for the understanding of an emoji, so we started the project to create a word cloud tool for plotting emojis while preserving all the initial properties of the emojis.

To start with, I looked into some website links related to plotting emojis as a picture in Python. From Dr. Feng's suggestion, I also looked up related works from Google Scholar and other sources, including Talkwalker, a brand-management company that also provides data analytics. Talkwalker, as well as Slido and Poll Everywhere, offers to plot emojis on a word cloud format, but their functions are either specific for collecting audiences' responses or not open source. Our EmojiCloud project is the first open-sourced word cloud tool for emojis.

On February 8, 2022, I found a git repository that contains all the 3577 emojis on Twitter. I started working with Twitter emojis to create the prototype of the EmojiCloud. I am also using sample emoji-frequency data from a previous project for determining the size of the emojis on the prototype. However, at the beginning of this project, I felt stressed because I could not think of a proper way to put all the emojis onto the canvas without overlapping. On the night of February 9, 2022, Dr. Feng and I talked for a long time on Zoom for cleaning up our idea and start with a simple 5x5 emoji plot. It is created by dividing the canvas into a 5x5 grid, locating the center of the grid, and plotting 25 emoji based on the distance of each center to the canvas center.

Dr. Feng provided me with his code on February 19, 2022. It plots all the emojis without occlusion and also the emojis are next to each other. After Dr. Feng revised the code to increase efficiency, I improved the prototype to add a 10-pixel distance between the emojis so the cloud looks more balanced.

On February 24, 2022, after trying to make the canvas into a flat rectangle, I started plotting all the emojis within a mask. Since I am using Twitter emojis, I used the Twitter logo as a mask for the prototype. After adapting the different formats of the mask image, I started working on adding a contour for the mask of the EmojiCloud around February 28, 2022. This is a process that tools me the most time when I roughly finished it on March 16, 2022. However, I am still trying to resolve the problem of some segments of the contour being thinner compared to others until May 11, 2022.

Starting March 21, 2022, I looked up several related works, including Talkwalker, Slido, and Poll Everywhere. I drafted some writings for the related works section of the paper. Before the final submission on April 10, 2022, I also proofread the draft and offered revising suggestions.