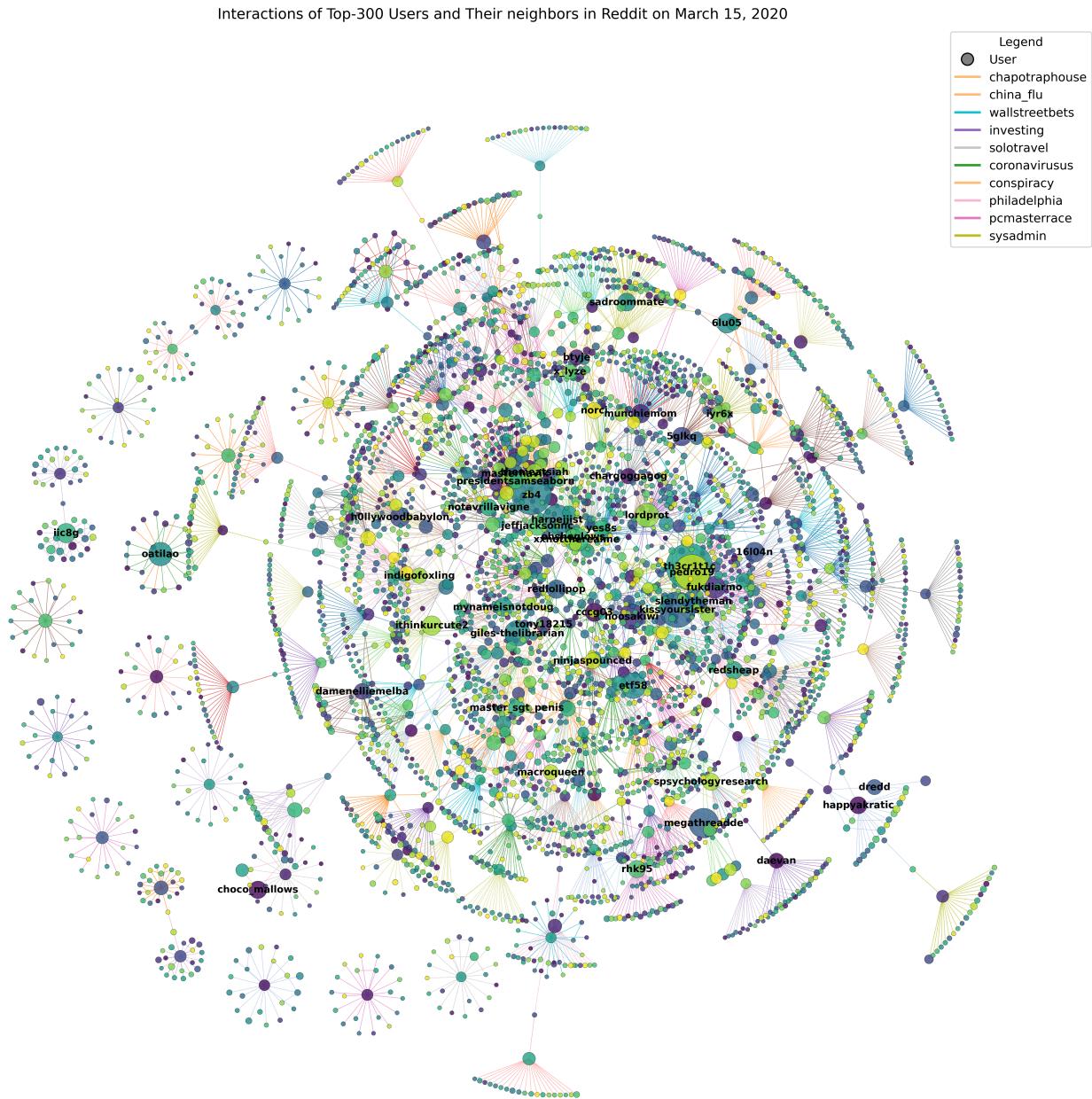


# INFSCI 2415 Final Project

---

Graph: **Interactions of Top-300 Users and Their neighbors in Reddit on March 15, 2020**



I use the Reddit comments and submissions on March 15th, 2020 to build the interaction graph. Only Top 300 users with highest degree and their neighbors are shown. Only top 50 users' name are shown.

- **Nodes:** Nodes represent users
- **Edges:** two nodes are connected when one user creates a comment to another user's submissions, or creates a comment to one user's comment.
- **Color of Edge:** The colors of the edges indicate the subreddit to which the comments belong.

## Highlights

- **Highly dense interaction core:** The center of the graph shows a tightly clustered core of highly active users, indicating intense cross-conversation among the top users.
- **Strong multi-subreddit overlap:** Many edges of different colors converge in the core area, showing that influential users frequently move across subreddits, facilitating cross-community information flow.
- **Peripheral “star-shaped” structures:** Surrounding the core are many small star-shaped or tree-like clusters, representing users who mainly interact with one or two highly active central users but do not interact much with each other.
- **Subreddit-specific micro-communities:** Some colors form distinct local clusters, suggesting that while top users bridge multiple communities, many ordinary users still remain subreddit-specific.

## Data and Method

I used Reddit comments and submissions data from March 15, 2020, focusing on reply interactions among users across COVID-related and high-activity subreddits. First, I removed all deleted posts (both comments and submissions), and posts by bot accounts and too-short posts. Then I filter submissions related COVID-19 by keywords, and keep these submissions and their all comments. Finally I built the network using comments per day to model users' interactions. The nodes represent users, and two nodes are connected when one user creates a comment to another user's submissions, or creates a comment to one user's comment.

In the visualization stage, First, I identified the top-300 most active users on that day based on total degree (sum of replies given and received). I extracted these top users and all their direct neighbors (users who replied to or were replied by that user) as the subgraph of the network. We then constructed the graph for visualization.

## Significance of the Presented Figure

This figure provides a clear picture of how online discussions during a major global event are structured across Reddit. By mapping interactions among the most active users and their communities, the visualization reveals how information spreads, how cross-subreddit communication occurs, and which users act as central hubs connecting multiple conversation spheres. Understanding these structural patterns is crucial for studying information diffusion, moderation dynamics, misinformation pathways, and the emergence of community polarization during crisis periods such as the COVID-19 pandemic.

## Reference

Code: [Github](#)