



reddit as a network

Search Reddit



adam maciaszek

27.9k karma

500

140 Introduction

Reddit differs significantly from other social networks in that it does not revolve around the idea of following friends or influencers. Instead, users are encouraged to engage with and contribute to various subreddits, each of which represents a distinct community centered around a specific topic. While previous research has suggested that social networks such as Facebook and Twitter have become increasingly isolated echo chambers, it remains unclear whether Reddit is similarly affected. Despite the existence of clearly defined communities on the platform, which often hold strong opinions, it is possible that the level of interconnectivity between subreddits has prevented the formation of isolated echo chambers.

This project aims to investigate the level of interconnectivity between subreddits and identify the subreddits that share the most users, forming larger clusters. To determine whether Reddit has always functioned in this way, we will utilize a dataset consisting of all the top posts of the top subreddits, scraped in 2013. Since then, Reddit has experienced substantial growth, with a five-fold increase in active monthly users, totaling 500 million. This study seeks to examine the impact of this growth on the interconnectivity of subreddits.

55 Comments Award Share Save ...

200 Methodology Scraping & Constructing the Graph

From the data collected in 2013 there were 2500 subreddits which the information on the top 1000 posts was collected using the python package praw. This included post text author link and post ID for further querying. Going through each post ID the all comments and replies to any comments the author and comment karma was recorded.

A log was created of each username that occurred their total comment karma of all comments on that particular subreddit. Using that information an edgelist was created with each edge being the common number of users which have left a comment in the top 1000 posts sample. The total number of users in the 2023 dataset was 67 million which is a good fraction considering the 500 million monthly users.

Due to several quarantined subreddits, only 2,416 subreddits were accessible for information retrieval during the data collection process for the top posts in 2023 of the same subreddits.

2013 Node Degree Distribution 2023 Node Degree Distribution

65 Comments Award Share Save ...

