Logan Carter
COM 411
Spring 2026

Network Analysis - Discord Server

Introduction

Discord servers are informal communication spaces where interaction happens constantly but not evenly. Some users are highly active and interact with many people, while others participate less frequently or only within smaller groups. Because of this, I chose a Discord server I am a part of as a useful setting for studying how communication organizes itself naturally in this environment.

My project focuses on analyzing user-to-user communication within a friend's Discord server through visualizations and analysis. My goal is to understand how users interact, who is most active, and how characteristics such as age, tenure, and interests relate to their communication behavior. Instead of only describing activity, I use network analysis to better understand the group.

To do this, I constructed a weighted network using survey data, full message history extraction, and AI/human analysis to create an edge and node attributes list. I used the final network visualizations and supporting attribute data to examine aspects of the server.

Network Theory

My project highlights three main network theory concepts: degree centrality, homophily, and core-periphery structure. I use these concepts to understand both how communication is distributed and why certain patterns may emerge.

Degree centrality is useful for measuring how active users are within a network. It is defined as the number of connections a node has, meaning how many other users it interacts with. In a communication network, users with higher degree centrality are more active and more connected to others. I use degree centrality to identify the most active or central individuals in the server (Golbeck, J.).

Homophily refers to the tendency for individuals to associate with others who are similar to them. This similarity can be based on characteristics such as age, interests, or behavior. Research shows that people are more likely to form connections with others who share similar attributes, because similarity can make communication easier and more predictable (Block, P., & Grund, T.). In my project, homophily helps frame whether users with similar interests, ages, or tenure appear together.

Core-periphery structure describes a common pattern in social networks where a group of highly connected nodes forms a central "core," while less connected nodes exist on the "periphery." Core nodes tend to be densely connected to each other and also connected to peripheral nodes, while peripheral nodes are less connected overall and often rely on core nodes for access to the broader network (Yang, J., Zhang, M., Shen, K.). I use this concept to examine whether communication in the server is centralized around a smaller group of active users.

Based on these concepts, I focused on the following research questions:

1. Who are the most active users in the Discord server?
2. What reasons do users report for being in the server?
3. What is the age distribution of users in the server?
4. How long have users been members of the server?
5. Does the network show evidence of a core-periphery structure?

Method and Data

I collected data from a friend's Discord server consisting of approximately 10 to 20 active members. To build the dataset, I used a combination of manual review, survey responses, and AI-assisted processing rather than relying on a single method.

First, I used a Google Chrome extension to extract all usernames from the Discord server. This gave me a complete roster of users and served as the foundation for the node attribute dataset. After that, I created a Google Form and distributed it to members of the server. The form collected information such as username, display name, role, activity level, age, tenure, and primary reason for being in the server. It also asked users to list who they communicate with and how often, which helped me create an initial version of the edge list.

Next, I used an external online tool that required my private Discord key to download the full message history from all server channels. This gave me raw interaction data, including message exchanges and reply behavior between users.

I then used AI models, ChatGPT and Claude, to create an initial version of the edge list by identifying direct reply interactions in the dataset. These direct replies were explicitly labeled in the data, which allowed the models to reliably assign weights between users who responded to one another. After generating this initial edge list, I manually reviewed the chat history and focused specifically on interactions that were not captured by the direct reply feature. This included conversational replies, name mentions, and responses that required interpreting context. I used this manual process to add new edges and increase the weights of existing connections where appropriate. This two-step approach allowed me to combine structured interaction data with more nuanced communication patterns that AI alone could not fully capture.

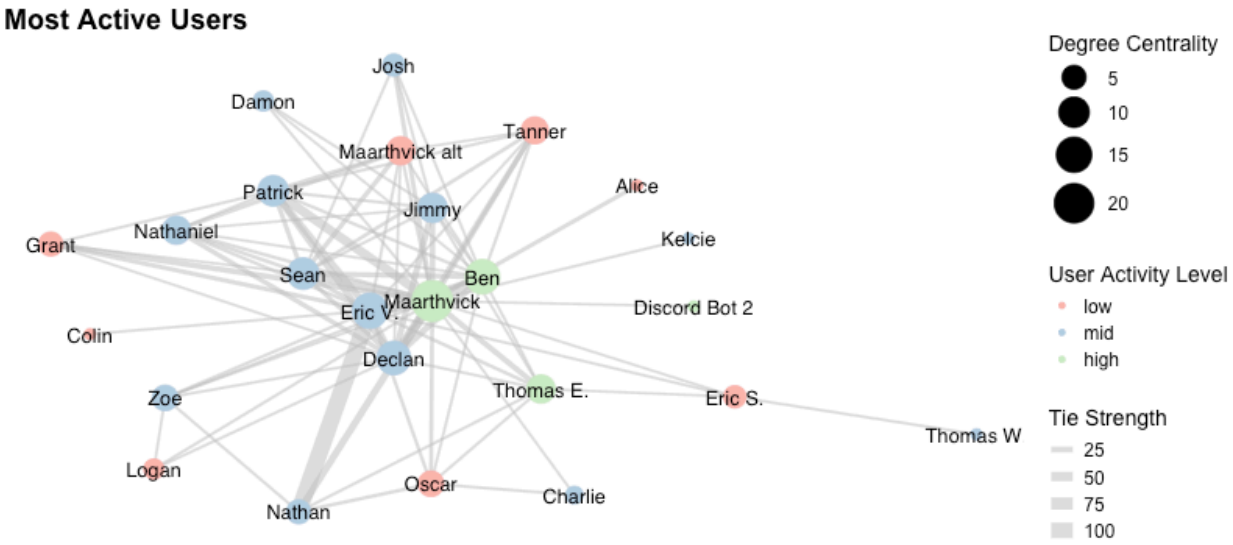
I determined the final edge weights by starting with the AI-generated edge list based on direct reply interactions. After that, I manually reviewed the chat history to identify additional interactions that the AI could not reliably capture, such as indirect conversational replies, name mentions, and responses based on context. I used this review to add new edges and increase the weights of existing edges when users interacted more frequently. I also considered the survey responses as supporting evidence for whom users reported communicating with.

To make sure all nodes were included, I also had the server owner complete the survey so that Discord bots and less active users were represented in the dataset. After finalizing the node attributes and edge list, I imported both Excel files into R and used igraph, tidygraph, and ggraph to construct and visualize the network.

There are several limitations to my approach. While AI was useful for identifying direct reply interactions, it could not reliably interpret more nuanced conversational behavior. Many messages do not explicitly indicate who they are directed toward, and AI models are not able to consistently recognize these implicit interactions. Because of this, the AI-generated edge list was incomplete and required manual correction. In addition, survey responses may contain bias or incomplete reporting, since users are estimating their own behavior. Finally, this dataset only includes communication within the Discord server and does not capture interactions that occur outside of it.

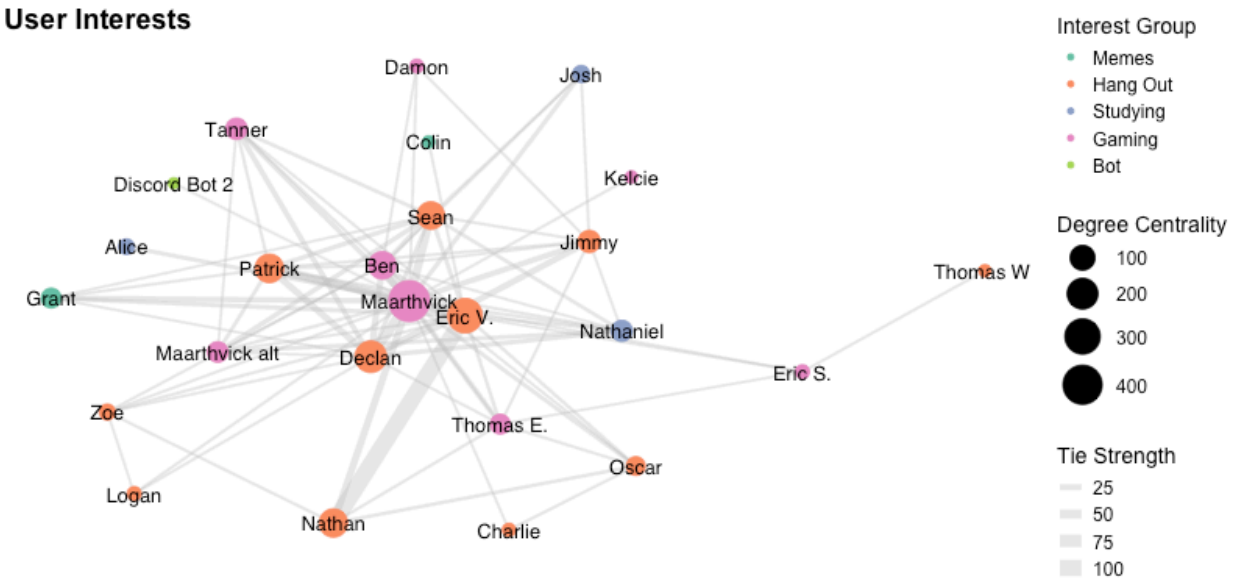
Analysis

My first research question focuses on identifying the most active users in the server. To answer this, I used degree centrality, which measures how many connections each user has in the network. Users with higher degree values are more active because they interact with more people and participate more frequently in communication. I represented this visually through node size, where larger nodes indicate higher activity. Node color in this visualization represents self-perceived activity level collected from the survey, which allows for comparison between reported and actual communication behavior. From this visualization, it is clear that a small number of users account for a large portion of the connections in the network, indicating that communication is not evenly distributed across members. It is also clear who the owner of the server is, as this user (Maarthvick) has the largest node size and the most ties.



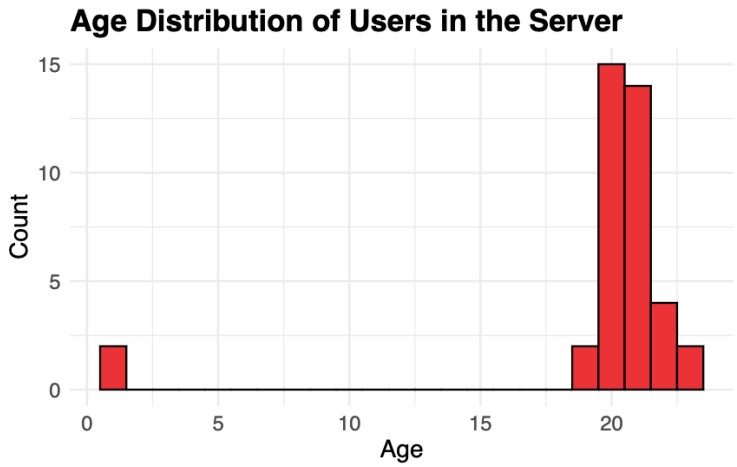
View Code

My second research question examines the reasons users report for being in the server. This information comes directly from the survey data collected in the node attribute table. By grouping users based on their selected interest categories, I use the visualization to show how motivations are distributed across the server and whether users with similar interests appear near each other in the network. From the visualization, there is some evidence of homophily, as the majority of users fell into two of the five interest groups (hanging out and gaming). However, grouping based on interests is not especially strong, likely because most users were only able to select one main interest group, which limits how much variation can be observed.



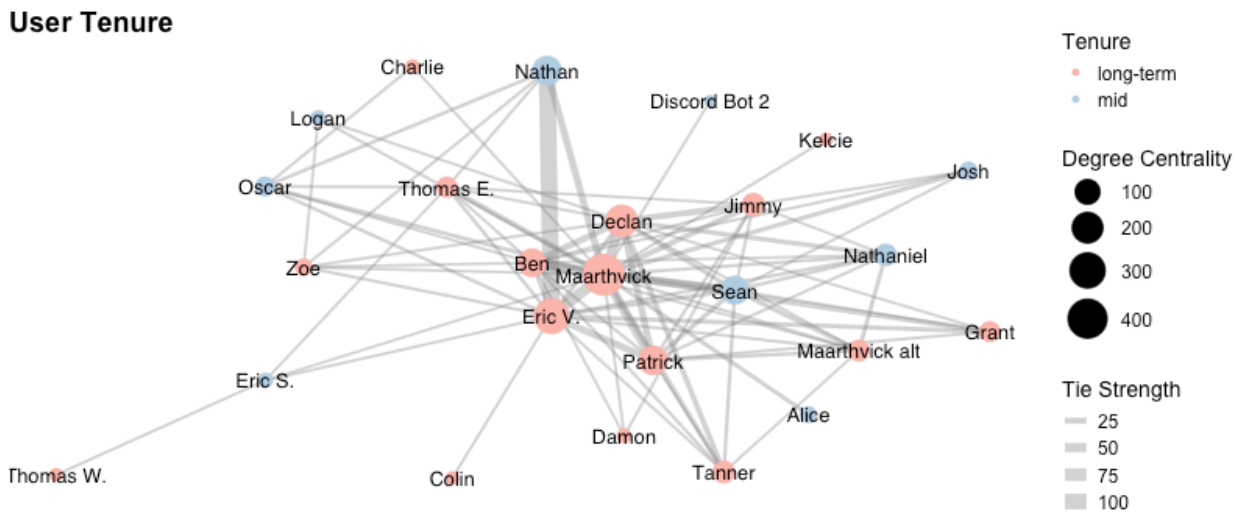
View Code

My third research question looks at the age distribution of users in the server. This is based on survey responses and provides a basic demographic overview of the group. I use the age visualization to show whether the server is mostly made up of users in a similar age range or whether ages are more spread out. Note: Discord bots were assigned an age of 1 for identification purposes, which slightly skews the lower end of the distribution but does not affect the overall pattern among real users. The distribution shows that most users fall within a relatively narrow age range, which is expected given that the server is made up of college students. This also helps explain why communication is dense, since users are in similar life stages and interact frequently.



View Code

My fourth research question focuses on how long users have been in the server. This helps provide context for understanding communication and connection patterns by showing whether longer-tenured users appear more integrated into the network compared to newer members. From the visualization, longer-tenured users appear more central in the network, which makes sense considering they would have more time to chat within the server.



View Code

My final research question examines whether the network shows evidence of a core-periphery structure. Based on the visualizations, the network does appear to show a core-periphery pattern. A smaller group of users appears more central and highly connected, while less connected users are positioned farther from the main structure. This suggests that communication is concentrated around a core group rather than evenly distributed across all users.

Discussion and Conclusion

My project shows that communication within a Discord server can be analyzed as a network to reveal the underlying structure. Even though the server appears informal, communication is not random. Instead, it is concentrated among a smaller group of highly connected users, which reinforces the core-periphery structure observed in the visualizations.

Degree centrality helped me answer the question of who is most active by identifying users with the highest number of connections. Homophily gave me a way to interpret patterns in user characteristics, especially in the overall makeup of the network. Core-periphery structure helped me think about whether communication is organized around a central group while other users remain more peripheral.

At the same time, my project has limitations. AI was useful for identifying messages using the reply feature in the exported data, but it could not reliably interpret every conversational reply or indirect interaction between users. Because of this, I manually reviewed parts of the chat history to add interactions that were not captured by the direct reply feature. Survey responses may also not fully reflect actual behavior, since users are estimating their own communication patterns. In

addition, my analysis only captures communication within the Discord server and does not include interactions that happen outside of it.

Despite these limitations, my project shows that network analysis is useful for understanding communication in online environments. Even in a small and informal Discord server, communication is not random. Structured patterns of interaction still appear, and these patterns can help explain who drives communication, how connected the group is, and how users organize themselves socially.

References

Anthropic. (2026, April). *Claude* [Large language model]. https://claude.ai/

Block, P., & Grund, T. (2014). Multidimensional homophily in friendship networks. *Network Science, 2*(2), 189–212. https://doi.org/10.1017/nws.2014.17

Golbeck, J. (2015). Analyzing networks. In *Introduction to social media investigation*. Elsevier. https://www.sciencedirect.com/science/article/pii/B9780128016565000214

moska. (2025, December 19). *dSaver - Export Discord user list to CSV* (Version 1.1.4) [Chrome extension]. *Chrome Web Store*. https://chromewebstore.google.com/detail/dsaver-export-discord-use/fpiffaomkpnhmldgjbcinoikplpjnpca?hl=en

OpenAI. (2026, April). *ChatGPT* [Large language model]. https://chatgpt.com/

Apify. (2026, April). *Discord Channel Messages Scraper* [Software]. https://apify.com/felt/discord-message-scraper

Yang, J., Zhang, M., Shen, K. N., Ju, X., & Guo, X. (2018). Structural correlation between communities and core-periphery structures in social networks: Evidence from Twitter data. *Expert Systems with Applications, 111*, 91–99. https://doi.org/10.1016/j.eswa.2017.12.042