

421 Project Proposal

The members of our group and what each person will work on.

Bassim Beshry - I am planning to analyze the network and see how users view counts are affected when streaming explicit content versus non explicit content. I also plan to work on the documentation of my findings for the final deliverable.

Zeyad Elganainy - I plan on analyzing how twitch affiliate status is related to a user streaming explicit vs non explicit content. What could being an affiliate entail in terms of the types of content a user can stream?

Jordan Pohr - I am planning to analyze the network to see the correlation between maturity rating of a streamer and how that affects their follower counts on the network. I also plan to help build the necessary plots needed to conduct a proper analysis of the network.

The problem we are trying to solve and the plan to achieve it.

We plan on analyzing the twitch network to identify Twitch users who stream explicit content based on their position in the follower network and interaction patterns. We hope to understand whether streaming explicit content affects variables such as a user's view count, their affiliate rank as a twitch creator and their follower count.

Through node classification (binary classification for explicit vs. non-explicit streamers), possibly using features like follower network density, follower characteristics, and streamer engagement patterns we hope to answer the questions below..

The research questions our project will answer.

We're aiming to explore how being a twitch affiliate affects a streamer's maturity rating And the correlation between the two (?)

Does maturity rating have an affect on a user's view count?

Does maturity rating have an affect on a user's follower count?

Where we plan to get the data from.

We are getting our data from the Stanford Network Analysis Project.

https://snap.stanford.edu/data/twitch_gamers.html

How we plan to define the nodes and edges.

Nodes: Streamers (Node sized based on view count, Node color based on maturity rating)

Edges: Followings

The metrics we plan to use and the analysis we plan to conduct.

Degree Centrality - This measures the number of direct connections a user has. Since this is an undirected graph, it's useful for determining how connected they are within the Twitch network. This metric helps us identify influencers and popular streamers, which will in turn help us determine their relation to view counts and explicit content.

Cliques - Cliques identify tightly-knit groups within the network, such as communities of streamers who follow each other. We will analyze cliques to see if there are communities around explicit or non-explicit streamers and how that affects their follower growth or view counts.

Modularity - We plan to use modularity to determine sub-communities of explicit or non-explicit content creators

Density - We plan to calculate density to determine how connected the entire network is. This will be used to answer whether explicit content streamers form dense subgroups within the larger Twitch network and how that might affect their engagement.

Eigenvector Centrality - We're hoping that by measuring eigenvector centrality we can analyze if a streamer who streams explicit content is followed by other influential users, which might amplify their reach.

Closeness Centrality - Users who have a high closeness centrality are likely to be quickly accessible by others (viewers/followers), which could help us analyze how that influences their affiliate status or follower count

Betweenness Centrality - Measuring betweenness will help us find streamers who act as bridges in the network, connecting different cliques or communities. It could give us some insight to understand how streamers who produce explicit content may connect otherwise disconnected user groups