

# Twitch Social Networks

## Background

Twitch specializes in live video game broadcasting, including coverage of sports events. It also provides music streams, original material, and "in real life" feeds. It is run by Amazon.com, Inc. subsidiary Twitch Interactive.

The primary goal is to evaluate and explore the network, learn about its size, structure, and subgroup measurements, topographical, interconnected measures. It also identifies the central actors and key players in the network and offers a way to market a product either directed towards the streaming community or the gaming community that follows a particular stream while maximizing your reach.

## Data

These datasets contain Twitch user-user networks of gamers who stream in a certain language. Nodes are the users themselves and the links are mutual friendships between them. These social networks were collected in May 2018.

C_Names <chr>	No_Nodes <int>	No_Edges <int>	NW_Diameter <chr>	NW_Avg_Dist <chr>	NW_Density <chr>	NW_Avg_Degree <chr>
DE	9498	153138	7	2.72157110579611	0.00339542673016202	32.2463676563487
ENGB	7126	35324	10	3.6776157289097	0.00139145506201653	9.91411731686781
ES	4648	59382	9	2.88319143955699	0.00549852272689393	25.5516351118761
FR	6549	112666	7	2.68099071395718	0.00525459454049372	34.4070850511528
PTBR	1912	31299	7	2.53237915700558	0.0171321505750675	32.7395397489539
RU	4385	37304	9	3.02109464082098	0.00388101440711117	17.0143671607754
ALL	34118	429113	NA	NA	NA	NA

7 rows | 1–7 of 8 columns

**Data Source:** <https://snap.stanford.edu/data/twitch-social-networks.html>

### Node Attributes:

- Name:** Name of the Node[Character]
- Days:** Days the streamer has been streaming[Numeric]
- Mature:** Is the Content Mature[True/False]
- Views:** Total Views for the Streamer[Numeric]
- Partner:** Partnership Status with Twitch[True/False]
- Twitch ID:** Twitch ID of the Streamer[Integer]

**Edges indicate Mutual Followership between Streamers.**

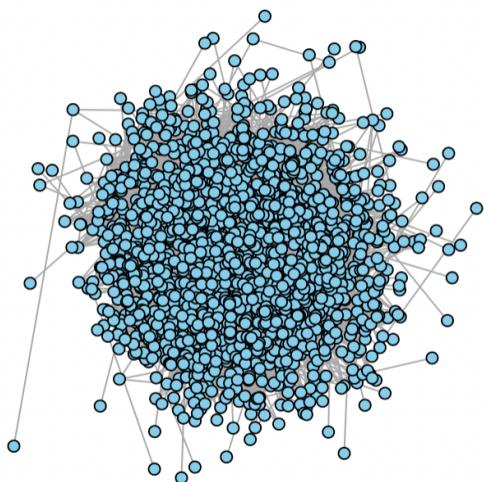
### **Filtering:**

For the marketing purpose we are focusing on Top 100 Streamers from Portuguese Brazilian speaking community for which the statistics are as follows:

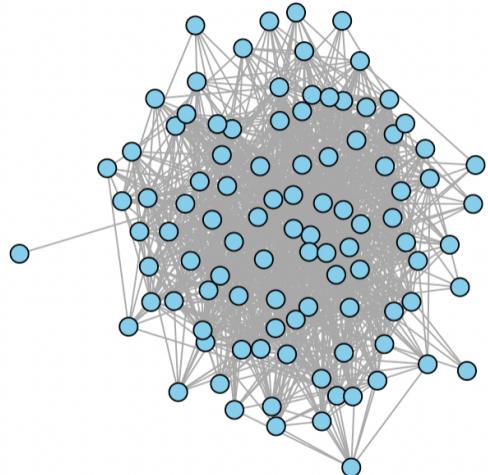
**Number of Nodes [Streamers]: 100**

**Number of Edges [Mutual Followership between streamers]: 1329**

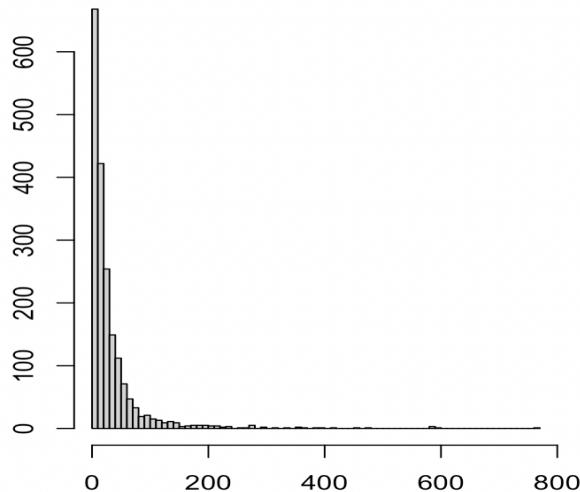
**Network Before Filtering**



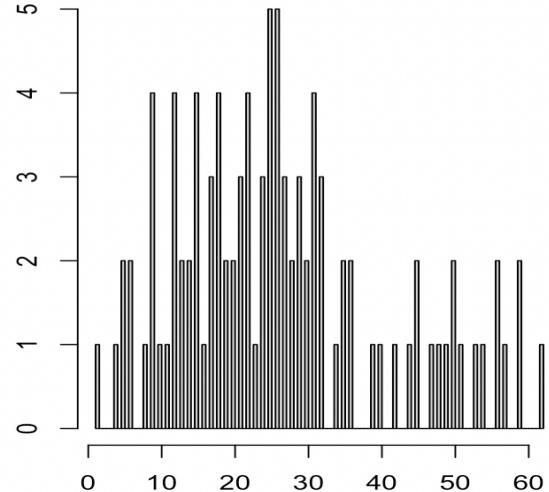
**Network After Filtering**



**Histogram of Degree Centrality  
Before Filtering**



**Histogram of Degree Centrality  
After Filtering**



## Example of Node Attributes:

	<b>name</b> <chr>	<b>days</b> <int>	<b>mature</b> <lgl>	<b>views</b> <int>	<b>partner</b> <lgl>	<b>twitch_id</b> <int>
ptbr_290	ptbr_290	1896	FALSE	116392808	TRUE	47071880
ptbr_1320	ptbr_1320	2490	FALSE	40659090	TRUE	26857029
ptbr_103	ptbr_103	2454	FALSE	40456102	TRUE	27680990
ptbr_1739	ptbr_1739	1781	FALSE	35175114	TRUE	52203144
ptbr_1660	ptbr_1660	2024	FALSE	33101623	TRUE	41783889
ptbr_127	ptbr_127	1592	FALSE	32672093	TRUE	63602976
ptbr_45	ptbr_45	1122	FALSE	18534438	TRUE	102346837
ptbr_1476	ptbr_1476	2195	FALSE	18475754	TRUE	36772976
ptbr_1259	ptbr_1259	2153	FALSE	17962557	TRUE	37825200
ptbr_305	ptbr_305	2225	FALSE	14861110	TRUE	36024998

## Analysis

Following image shows the values of Basic Topographical, Centralization & Interconnectedness Measures of the Twitch Social Network:

<b>Metric</b> <chr>	<b>Values</b> <dbl>	<b>Metric</b> <chr>	<b>Values</b> <dbl>
Size	1.000000e+02	Density	2.684848e-01
No_Edges	1.329000e+03	Average Degree	2.658000e+01
Diameter	3.000000e+00	Cohesion	1.000000e+00
Average Distance	1.767273e+00	Compactness	5.782913e-03
Degree Centrality	3.577778e-01	Clustering Coefficient	4.518495e-01
SD Degree Centrality	1.452840e+01		
Closeness Centrality	3.048187e-01		
SD Closeness Centrality	6.184647e-02		
Betweenness Centrality	3.982346e-02		
SD Betweenness Centrality	5.037615e+01		

## Topographical Measures:

### Questions:

How many actors are there in the network to be considered?

What is the number of intermediate acquaintances required to maximize the product reach?

How fast/easily the marketing will reach the targeted audience?

**Network Size:** Total number of streamers being considered for this marketing study. This subset has been taken based on the number of views the streamer has which is

assumed as his/her/their popularity in the community. Currently we are only considering Top 100 Portuguese Brazilian Streamers based on their viewership.

**Network Diameter:** Diameter of the network is the longest of all the shortest paths (geodesics) that traverse the network. Here the diameter of 3 indicates the maximum number of intermediate acquaintances a streamer would require if we want the marketing to reach the entire network.

**Average Distance:** Average Distance will tell us how quickly any information may spread within a network. Here an average distance of 1.7672 indicates that a second connection will cover most of the network maximizing our product reach.

## Interconnectedness Measures:

### Questions:

How well are the streamers interconnected so that marketing campaign can reach the maximum audience in the shortest time?

**Density:** Density indicates how interconnected the network is. Here the density of 0.2685 indicates that the network has 26.85% of ties/connections compared to maximum possible ties. Thus being densely connected / highly interconnected.

**Average Degree:** Average degree is the sum of each actor's degree centrality divided by the number of actors in the network. Indicating that a streamer on average is connected to 26.58+ other streamers via mutual followership.

**Cohesion:** Network cohesion is the ratio of all the pairs of actors that are directly or indirectly connected to the total number of actor pairs. Here the cohesion of 1 indicates that all the actors in the network can reach every other actor in the network directly or indirectly.

## Sub-Group Measures:

### Questions:

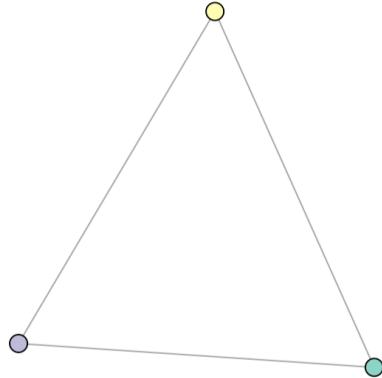
Is there a possibility to create customized campaigns for different groups to increase the engagement/reach within each group?

**Community Detection:** Tried to find possible communities using fastgreedy, spinglass algorithms. Fastgreedy & spinglass resulted in almost similar communities. Community

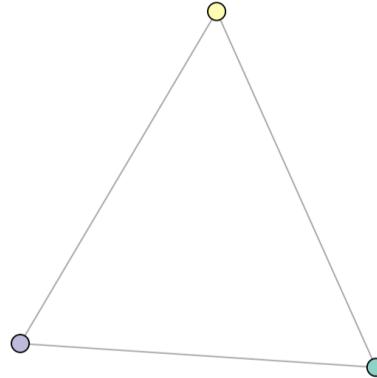
detection is well visualized with spinglass in 3D space compared to 2D representation. 3 communities are detected (out of which two are major for fastgreedy) & campaigns can be customized according to the users in each community.

## Simplified Communities

**FastGreedy Community**

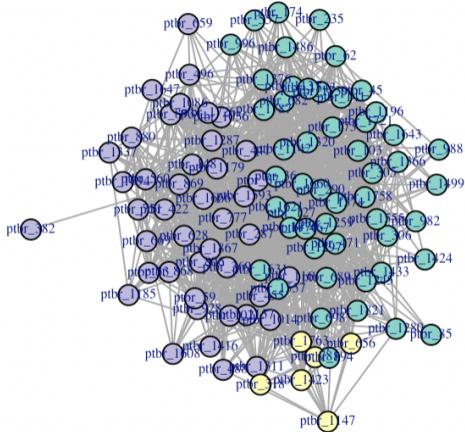


**Spinglass Community**

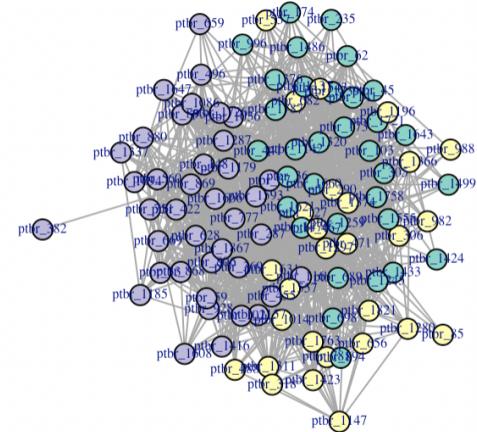


**Communities generated using FastGreedy and SpinGlass algorithm:**

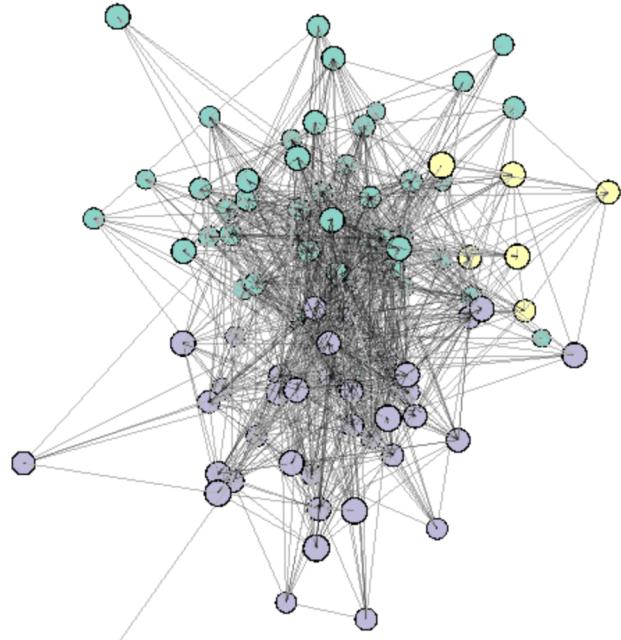
**Fast Greedy Community**



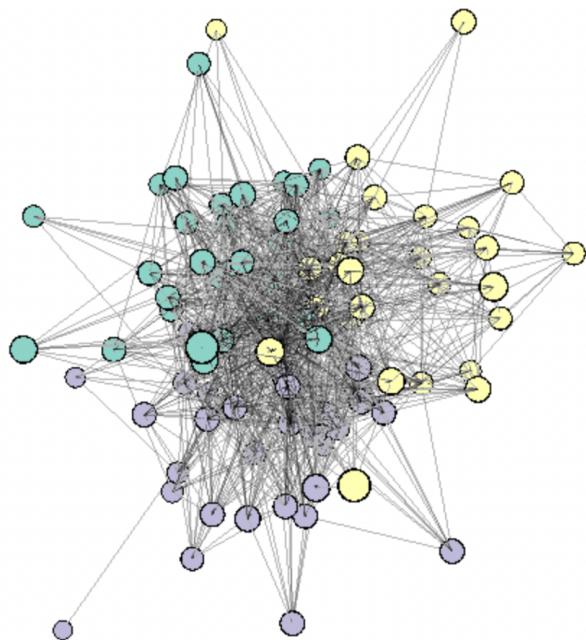
**SpinGlass Community**



**ThreeJS Representation of FastGreedy communities:**



**ThreeJS Representation of Spinglass communities:**



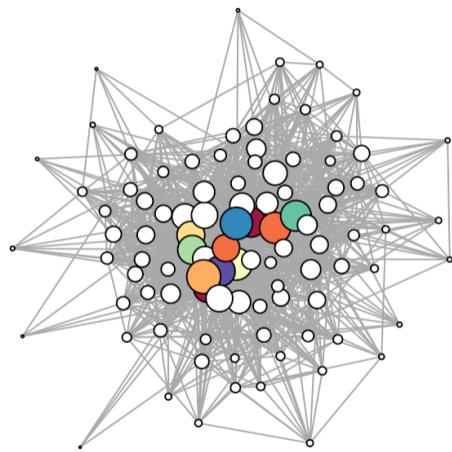
## Centrality Measures:

### Questions:

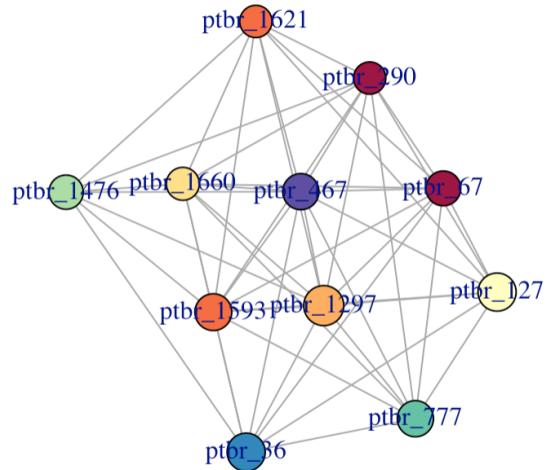
Which of these streamers can reach other streamers the most based on centrality measures?

**Degree Centrality:** The count of an actor's ties. Degree centrality can assess one of numerous aspects of a network actor. In the network, the streamer that has the highest degree of centrality can be a good target for marketing purposes. Because this streamer interacts with other streamers the most.

Degree Centrality

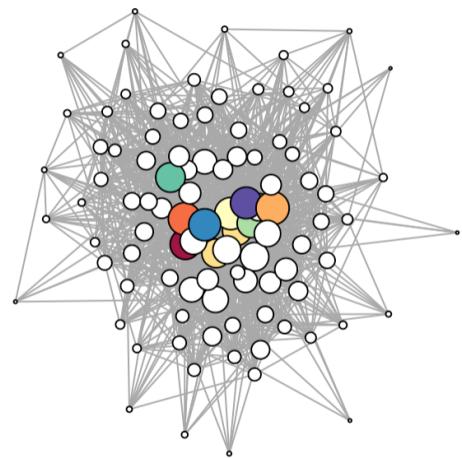


Degree Centrality for Top 11

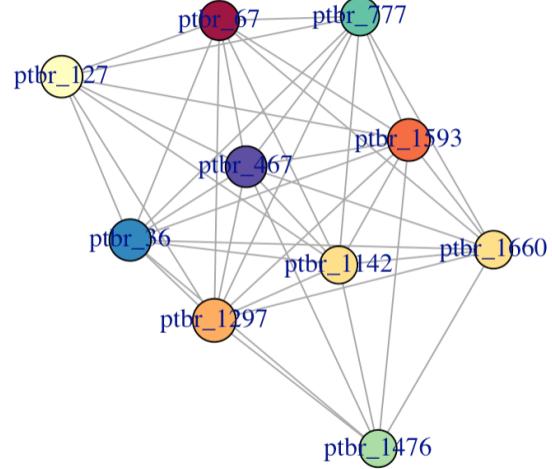


**Eigenvector Centrality:** Frequency-based centrality metric that measures the indirect power of an actor or their potential social capital within a network. In the network. The streamer with highest Eigenvector centrality will make a good target for marketing purposes because of the connection ties that the streamer has are the most powerful among all streamers.

**EigenVector Centrality**

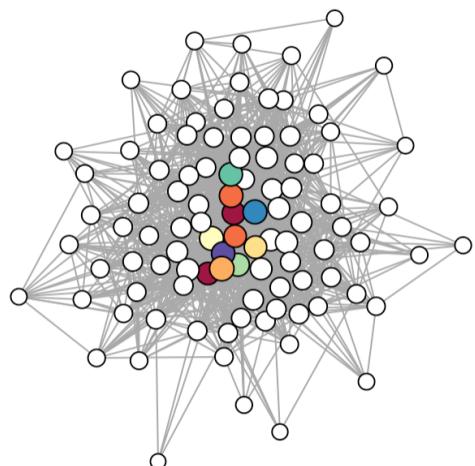


**EigenVector Centrality for Top 10**

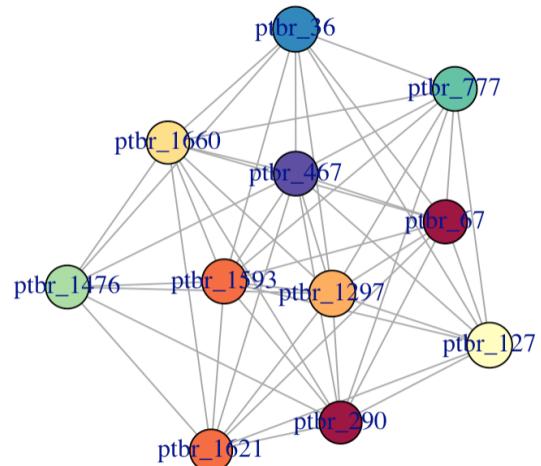


**Closeness Centrality:** The average geodesic distance from an actor to every other actor in the network. It adds up all of the shortest paths between each player in the network. The streamer with high closeness centrality can make a good target because of the high average shortest path.

**Closeness Centrality**

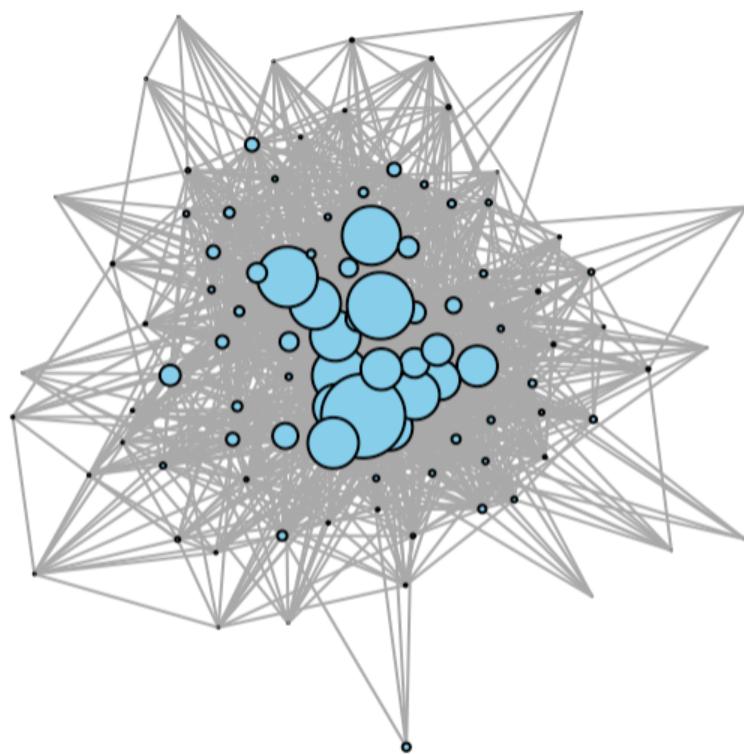


**Closeness Centrality for Top 11**



**Betweenness Centrality:** How often does an actor lie on the shortest paths between all pairs of actors. It is assumed that an actor has potential influence over any two other actors when it is on the shortest path between them. The streamer with the highest score of betweenness centrality can also be the target of the marketing purposes because this streamer lies on the shortest paths between streamers the most.

## Betweenness Centrality



*Here the nodes are sized according to their betweenness centrality for visual purpose.*

## Brokerage in the Network

### Questions:

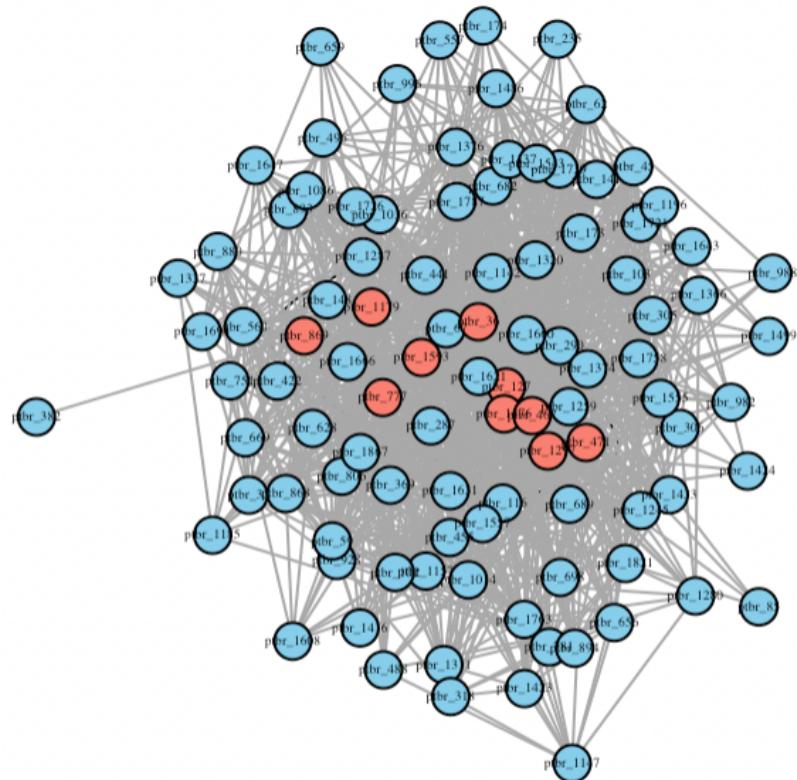
Does removal of a few nodes affect the entire network at all?

### Key Players:

Set of actors whose removal would either completely disconnect a network or maximally fragment it. *Key players are highlighted in the salmon color.* In our case, removal of key players wouldn't disconnect the entire network because of its dense connections. However, removal of these key players would decrease the efficiency of maximizing the reach in the network.

Key players in the network: "ptbr\_127" "ptbr\_1297" "ptbr\_869" "ptbr\_1179" "ptbr\_1593" "ptbr\_36" "ptbr\_777" "ptbr\_467" "ptbr\_471"

## Network w/ Key Players



## **Regular Equivalence:**

**Questions:** Can we prioritize the communities based on positional roles?

The main goal of regular equivalence is identifying roles based on regularities in patterns of relationships seen in a network. This can be used to create customized campaigns for different groups based on levels of connectedness. For example, from the below graph, it can be inferred that targeting yellow colored nodes will increase the engagement/reach more compared to blue colored nodes followed by green colored nodes.

## **Twitch PTBR top 100 Streamers REGE plot**

