Indian Institute of Technology Delhi

ELL 880: Social Network Analysis – Fall 2023
Project

Presented by:

Ranjan Kumar Singha [2023EET2192]

Sagar Karar [2023EET2755]

Course Instructor: *Sougata Mukherjea*

Contents:

Introduction & Motivation :

Online poker is a lucrative business for hosting sites and a passion for many enthusiasts. We are interested in understanding the dynamics of how the elite players (i.e. the professionals) interact with others in the network, including how they select which opponents to play with and where their winnings come from Also a major focus is been on the result of the game play between two players who have never played against each other and predict the outcome.

Problem Statement :

To determine if any concrete differences exist between how poker players of different skill levels and styles connect to the graph. Also, determine key summary statistics that predict how money will flow on the network of players and which types of players they fare better or worse against.

Approach and Methodology

We will create an edge creation model for an online poker network in which nodes seek inward edges (i.e. "profit") and attempt to avoid outward edges (i.e. "losses"). It includes a summary of standard network statistics (degree distribution, clustering coefficient, PageRank) along with a probabilistic model for network evolution based on preferential attachment.

The player interactions graph is shown below:

Player Interaction Graph

Metrics used for Analysis of online poker network

Net-Flow – The netflow is calculated as the difference between the netinflow(profit) and netoutflow (loss) of a graph .If the net flow is positive or higher than threshold value then it determines that the player or the node has gained net profit as a outcome from the game it has played.

Netflow = Netoutflow - Netinflow

The Clusters being formed as shown below are the positive flow and negative flow clusters for the players

Player Interaction Graph with Clusters

Green: Positive Inflow Cluster

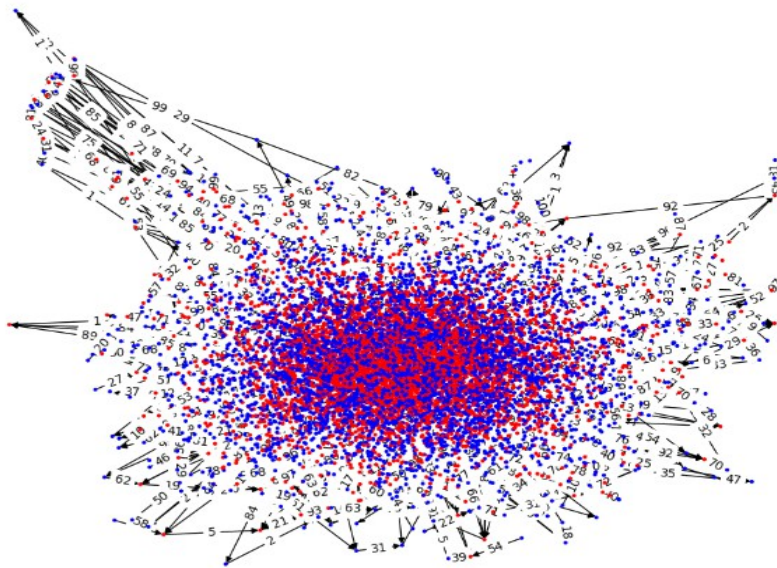Orange: Negative Inflow Cluster

Average degree of node- The degree of a node Is an important metric in the context of the no. of games played by a player .If the player has played more games then its degree should be higher and more than the average degree of the network.

Average degree of a node is calculated as sum of the degree of nodes divided by the total no. of nodes which can be used a threshold for the degree metric such that if a player has more degree than average degree then it should belong to the high degree cluster else it belongs to the low degree cluster .

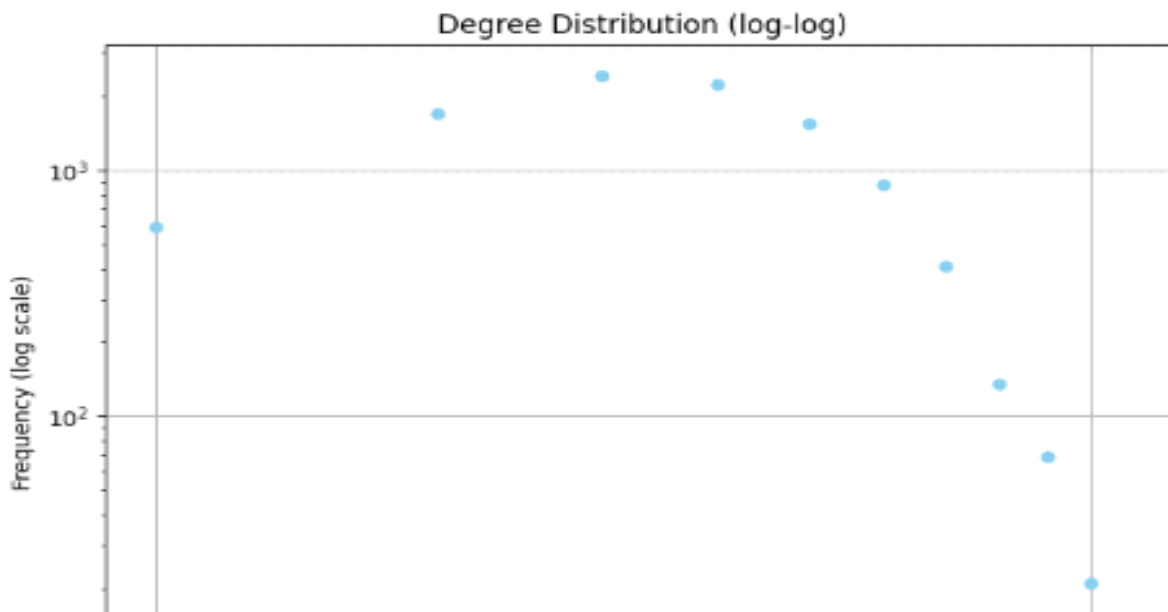The graph below shows the two clusters based on the degree metric :

Degree Distribution of the network is shown below in the given histogram:
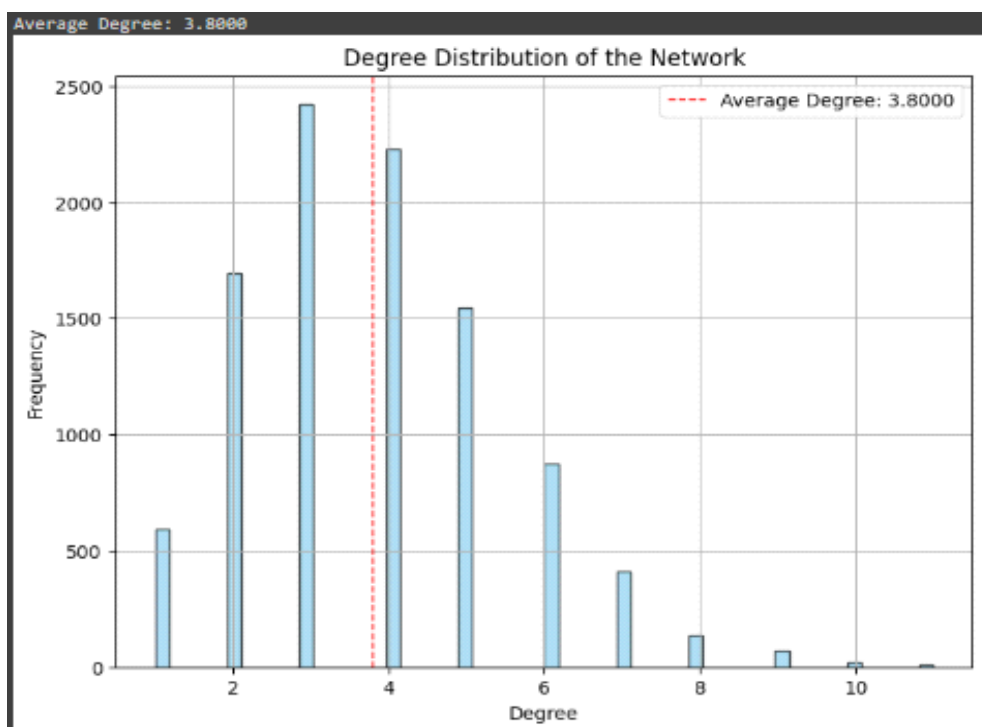
The degree distribution follows power law in the log log scale.

Degree Distribution (log-log)

Pagerank - the page rank of the nodes determines the pro-ness of a player i.e a player with higher page rank value have high percentage of match wins . Indirectly we can say higher the page rank better the player is and is tough to win against.

The page rank for a node is calculated as:

The PageRank Formula of a page v

$$Rank(v) = \frac{d}{N} + (1-d) \left(\sum_{u_i} \frac{Rank(u_i)}{outlink(u_i)} + \sum_{u_j} \frac{Rank(u_j)}{N} \right)$$

- d: Jump factor.
- N: Number of webpages
- u_i: Pages with links to v.
- u_j: Pages without outlinks.
- $Rank(u_i)$: The rank of the page u_i in the previous iteration.
- $outlinks(u_i)$: The number of pages u_i is pointing to.

The page rank based cluster is shown below:

Player Interaction Graph with PageRank-based Clusters
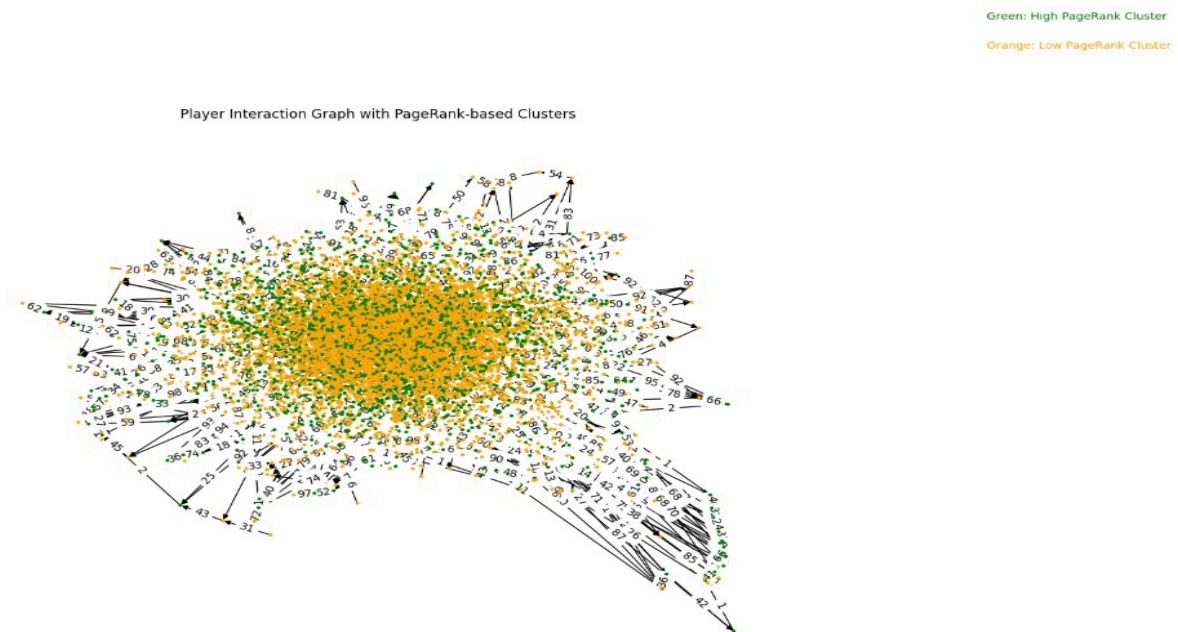
Hub Creation and authority score: The hub is created as the intersection of the postive outflow cluster, high degree cluster, high pagerank cluster which is termed as the hub of high quality players

Authority score importance – The authority score of a node can be used to determine the number of players of the high quality players cluster that a particular player has defeated .Thus it is used as metric for the node to determine whether a player has beaten any pro player or not.

The authority score of a node is calculated as the no of indegree edges from the hub .

The hub of high quality players is shown below:

High-Quality Players in Player Interaction Graph
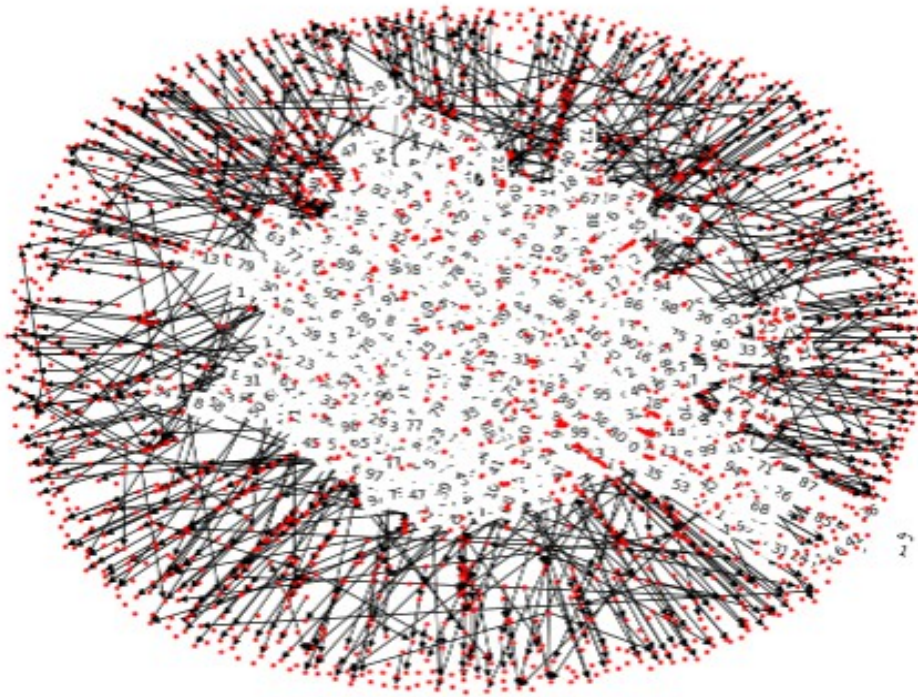
Results:

Any particular node can be analysed in the network by nodes of the network with whom the node has not played with .The screenshot below shows the nodes with which node number 6 has not played with :


```
5000
5001
5002
5003
5004
5005
5006
5007
5008
5009
5010
5011
5012
5013
5014
```

The main objective is to calculate the outcome of a node between two players who have not played earlier in the network.

The likelihood winner is calculated as:

```python
if score_A > score_B:
    print(f"Player {player_A} is likely to win.")
elif score_B > score_A:
    print(f"Player {player_B} is likely to win.")
else:
    print("It's a close match, and the outcome is uncertain.")
```

```
Enter the Player 1: 5
Enter the Player 2: 4536
Score of player 1:  4
Score of player 2:  5
Player 4536 is likely to win.
```

```
Enter the Player 1: 203
Enter the Player 2: 405
Score of player 1:  4
Score of player 2:  4
It's a close match, and the outcome is uncertain.
```

Conclusions and Future Work :

Transactional Interactions: Understanding the flow of funds between players sheds light on the network's financial dynamics and player behavior.

Network Structure: Identification of influential players, hubs, and patterns within the network contributes to comprehending its topology and potential vulnerabilities.

Future Scope: Implement machine learning algorithms for predictive analysis, anomaly detection

Reference:

CS 322: Social and Information Network Analysis Professor Jure Leskovec, Autumn 2009.

Analysis of Online Poker Network : Graham Yennie, Michael Jermann, Kartik Trasi, Stanford University, December 11, 2016

J. Saramaki, M. Kivela, J.-P. Onnela, K. Kaski, and J. Kertesz. Generalizations of the clustering coefficient to weighted complex networks. Physical Review, 2007.

Zhang and S. Horvath. A general framework for weighted gene co-expression network analysis.