Analysis of Online Poker Network

Graham Yennie, Michael Jermann, Kartik Trasi
Stanford University, December 11, 2016

Introduction

Problem Statement

Online poker, despite its rocky legal status in the US, 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 tables/opponents to play with and where their winnings come from.

Goal

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.

Related Work

Eng, Sharma 2009[3]
Summary: The authors of this paper developed 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.
Relevance: Despite the goals of this paper differing from ours (they were interested in the microevolution of individual nodes, whereas we are looking to identify "segments" of players and their optimal play styles), their formulation process of an online poker network helped us identify potential pitfalls.

Shah, Beutel, et. all 2015[4]
Summary: This paper provided an overview of anomaly detection algorithms across all types of graphs (unattributed, edge attributes, node attributes, etc.). It then proposes and implements a generalized anomaly detection algorithm for all types called EDGECENTRIC. The algorithm creates node clusters based on top-line attributes and identifies the nodes that vary the most from these clusters (by "minimum description length") as anomalies. The paper implements and analyzes the performance on 3 datasets from different industries (one of which they had ground-truth data) and were able to verify its high precision.
Relevance: Clustering nodes by top-line attributes and analyzing their interactions with the rest of the graph was the way in which we identified different "playing styles". The foundation of the EDGECENTRIC algorithm outlined served as an inspiration for our cluster-based network model.

Leskovec, Huttenlocher, et. al 2010[5]
Summary: This paper presents a theory of status that explains the observed signed (positive and negative) relationships between nodes which may provide an explanation into the underlying social phenomena. The paper claims that the phenomena seen in social networks where the relationships are defined by positive and negative relationships between entities cannot be fully explained by classical theory of structural balance hence the motivation to develop this alternate theory.
Relevance: This paper compares and evaluates the classical theory of structural balance and the proposed theory of signed networks using status theory using empirical data from social networks. It was an excellent resource when determining how to represent the positive and negative outflows between nodes in an online poker graph.

Lin , Jheng, et al. 2012[6]
Summary: The authors analyze a network of auction account holders that rank each other's profile. The goal is to find suspicious activity among users in a collusive structure. In this structure a fraudulent account is accompanied by many accomplice accounts that rank the fraud highly after small sham transactions. The authors modified a PageRank algorithm to weight activity that is more suspicious (i.e. opening an account and ranking another account in the same day is suspicious). After creating a modified PageRank ranking of suspicion, the authors applied an ANFIS algorithm to infer whether activity was collusive and fraudulent on a known test data set. Their performance was measured by an F-measure taking into account true positives, false positives, true negatives, and false negatives.
Relevance: The online poker network used was similar in structure to the auction site. While our goal was not to detect collusion, we did use PageRank similarly to indicate player "ability" in the weighted, directed graph; therefore, the methodology outlined here was extremely useful.

Methodology

Dataset

4.69 Million poker hand-history logs were cleaned into a game-by-game summary table, play-by-play summary table, and overall statistics of the players were computed. Raw data log format with player names and site identification hidden for privacy. The play-by-play table has 24.5M entries, the game-by-game table has 4.69M entries.

Terminology

VPIP	Voluntarily Put Into Pot – This stat measures the fraction of hands that a player engages in a hand, excluding when s/he is in a blind ante position and simply folds.
PFR	Pre-flop Raise – This stat measures the fraction of hands that a player raises during the first round of betting before the flop.
WTSD	Went to Showdown – This stat measures the fraction of hands that a player plays all the way to the end of the river and a showdown occurs.

Winrate	This stat measures the expected rate of profit a player has and is measured in terms of big-blinds per 100 hands (bb/100).

Classification of Playing Styles

A breadth of training books have been written on the subject of playing poker against specific opponents. Most of these texts focus on high-stakes poker games and would not stand up to the amateur-dominated graph in our dataset. On the topic of small-stakes cash games, Jonathan Little summarizes 6 distinct playing styles[1]:

Loose Passive (LP)	A loose passive player tends to call many hands preflop and raises much less frequently. These players tend to call more after the flop, unaware or untrusting that they are beat with marginal hand strengths. Because of this, a good counter-strategy is to bet against these players when you have value and avoid bluffing them and get out of their way when they show strength.
Tight Passive (TP)	These players will have a higher VPIP than PFR, but also have low VPIP. They will often spin "traps" with very strong hands. A counter-strategy is to exercise caution anytime these players are engaged in a hand with you.
Tight Aggressive (TAG) Straightforward	These players are much like the Tight Passive players, but will raise with more hands preflop. Their skill of disguising strength will diminish as the hands go on and will only tend to go to showdown with strong hands. In short, these players will enter more pots and fold too often. A good counter strategy for these players is to play aggressively late in the hand to force them to fold and fold when they display strength.
Tight Aggressive (TAG) Good	These players are much like the Tight Aggressive players above, except they will make more adjustments to their opponents. They are aware when they are in a hand with players that fold or call too much and adjust their strategy accordingly. A counter strategy against these players is to constantly act aggressively against them, but if there are easier players to play against, avoidance is advised.
Maniacal Loose Aggressive (Fish)	These players play so many hands and apply so much pressure, it is difficult to know their hand strength. A counter strategy for these players is to call submissively with strong hands to induce more bluffs from them and to bluff them intelligently.
Calling Stations (CS)	These players will look a lot like Loose Passive players, but will rarely fold any hand they determine to have value. These players are easy to beat by betting against them and rarely bluffing them

Game Selection

In another cash game text by Jonathan Little, game selection is a key skill for long-term success[2]. The logic implies that if you are a great player, but you only select games with better players, you will lose in the long-run. Apart from selecting easy games, selecting seats at a table where you have predictable players acting after your turn and unpredictable players acting before you is the most profitable strategy. In a weighted-edge graph, this effect will be hard to observe at a granular level, but will will focus on observing the tendencies of different playing styles in how they connect to the graph over many hands.

Graph Theory and Playing Styles

While top line statistics can aid a player at the tables, it does not tell a complete story about their opponents' overall strategy. Many players online use a head's up display (HUD) that tracks their opponents' statistics and helps them make decisions. We will model the online poker network where players represent nodes and edges are games that players have been in together. There are two versions of this we will look at: (1) an undirected network and (2) a directed network in which the edge attribute is the net monetary outflow from one player to another. We will calculate PageRank and Betweenness centrality measures and use our results from centrality measures to enable our ability to distinguish between player styles that align well with expert poker strategy guides, like the above description of playing styles. Playing styles will be detected using unsupervised learning algorithms using both the top-line statistics and those statistics with centrality measures included. These detected playing styles will be analyzed against each other to determine if the playing styles: qualitatively align with the style descriptions and counter strategies and quantitatively to determine significance of the interactions in terms of fitting winrates against the style interactions.

Betweenness Centrality

The betweenness centrality was calculated on an undirected graph as follows:

$C_B(v) = \sum_{s \neq v \neq t} \frac{\sigma_{st}(v)}{\sigma_{st}}$ where σ_{st} is the number of shortest paths from node s to node t and $\sigma_{st}(v)$ is the number of shortest paths from node s to node t passing through v.

PageRank

We calculated the PageRank of each node on a directed, weighted graph as follows:

$r_j = \sum_{i \to j} \beta \frac{r_i w_{ij}}{w_i} + (1-\beta)\frac{1}{N}$ where r_i is the PageRank of node i, $1 \leq i \leq N$, w_{ij} is the weight of edge from node i to j, w_i is the sum of weights of out edges of node i, and β is the teleportation parameter (0.85).

We used the PageRank algorithm as implemented in Stanford's Snap library. Each node is initially assigned a default PageRank of $\frac{1}{N}$ and continually updated using equation above until the values converge.

Results and Findings

Network Analysis

To gauge whether players of different skill levels select games differently, we started by calculating centrality of an undirected network in which nodes represented players and an edge means the players have been in a game together. Centrality was not correlated with net winnings.

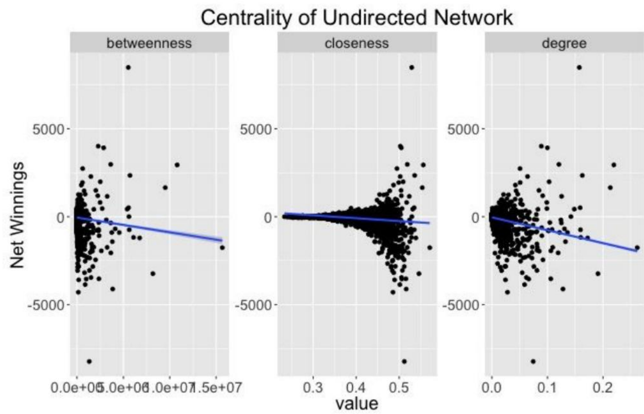


Figure 1. Centrality Metrics vs. Net Winnings for Undirected Network

We also bucketed users by their net winnings. The average clustering coefficient of these users is actually slightly lower - but much higher than the worst players.

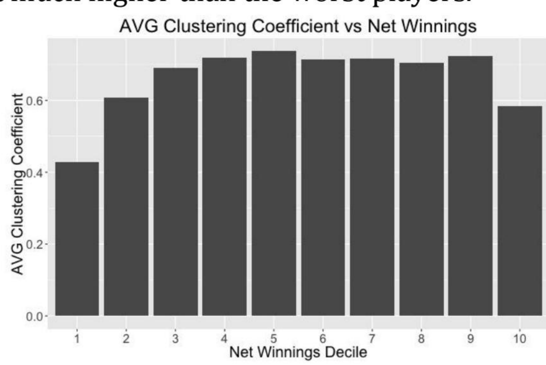


Figure 2. Clustering Coefficient vs. Net Winnings Decile for Undirected Network

This seems to indicate that the top players are playing more within the same community. In order to understand the actual flows of money within the network, we considered the dataset as a directed graph with the net outflow of money as the edge attribute. The transactions between nodes were aggregated into a single edge transaction that represents funds from source node to destination node. Weights are all represented as a positive number and the direction representing the flow of funds. Using PageRank on this network as a measure of centrality is better correlated to net winnings than on the undirected network; however, total number of wins is a much better fit.

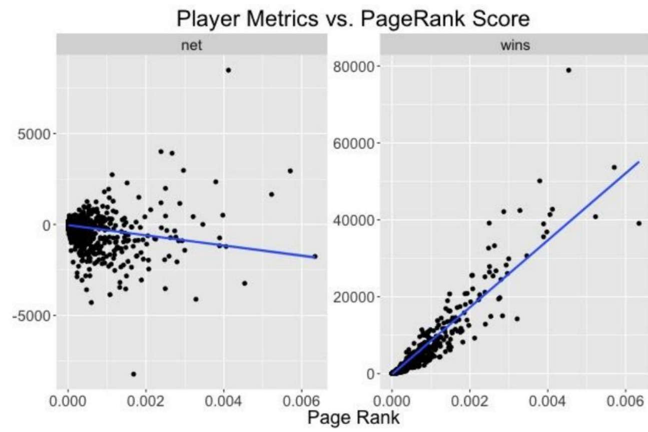


Figure 3. Weighted PageRank vs. Net Winnings and Win Count

To benchmark the winrate we simulated Erdos-Renyi and Preferential Attachment models with the same number of nodes and edges as the empirical network. The payoffs per hand for the synthetic simulations were drawn randomly from the empirical dataset without replacement. From the results (figure below) we observe that the empirical data is even more distributed than the preferential attachment model. We also see much higher count of players with low winrate (positive and negative) than in the Erdos Renyi and Preferential Attachment models.

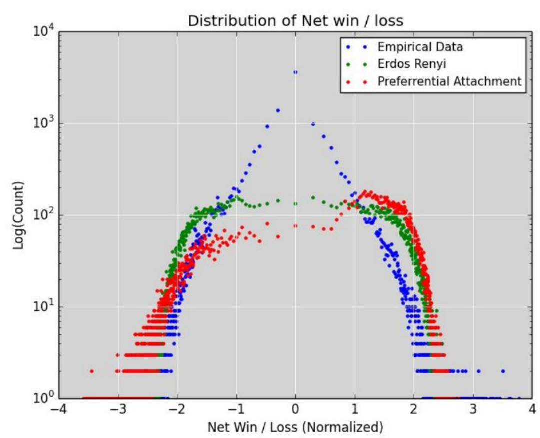


Figure 4: Empirical Data and simulated Erdos Renyi and Preferential Attachment

Unsupervised Clustering Analysis – Detecting Playing Styles

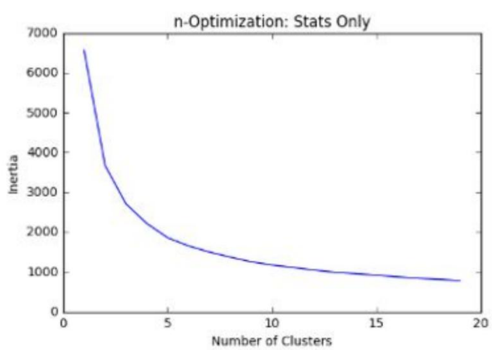
Three models were constructed to detect playing styles in our graph using K-means clustering analysis:

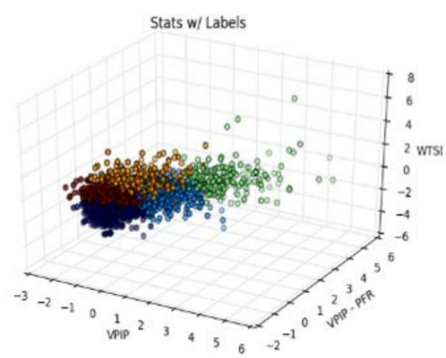
1. Top-Line Statistics Only: VPIP, VPIP-PFR, WTSD
2. Top-Line Statistics with PageRank Scores (logarithmic transformation)
3. Top-Line Statistics with Betweenness Scores (logarithmic transformation)

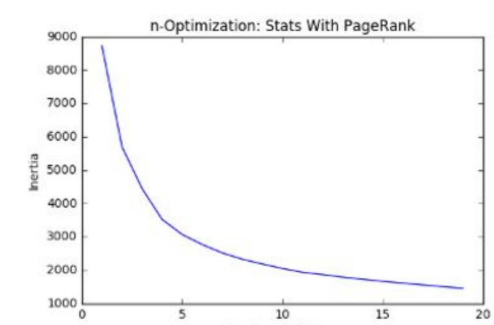
In each model, only players with 1000 hands or more were considered. The elbow method was used to determine a reasonable number of clusters to include in the algorithm, selecting a number of clusters based on the rate of change of inertia with respect to the previous number of clusters (forward selection). The figure below shows the inertia curve of the respective models and the resulting clusters in a selected 3D space. In each case, 5 clusters were selected as the parameter. Player styles were assigned qualitatively using the intuition from the player style descriptions above.

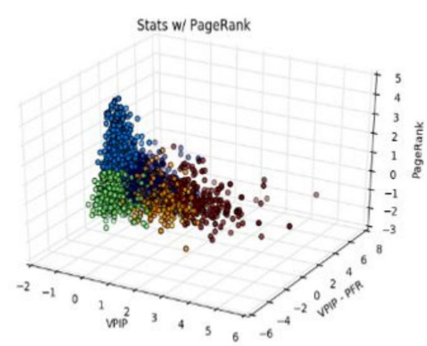
Stats Only								
Cluster	Winrate	VPIP	Agg Delta	PFR	WTSD	Betweenness	PageRank	Qualitative
0	-4.39	0.237	0.069	0.168	0.284	20893	5.6630E-05	Tight Aggressive
1	-7.86	0.507	0.342	0.165	0.283	6356	3.5650E-05	Fish
2	-3.12	0.248	0.068	0.180	0.285	29693	5.6780E-05	Tight Aggressive
3	-4.57	0.365	0.205	0.160	0.283	21341	5.7350E-05	Loose Passive
4	0.74	0.321	0.134	0.187	0.283	22353	4.6160E-05	Loose Aggressive
Stats with PageRank								
Cluster	Winrate	VPIP	Agg Delta	PFR	WTSD	Betweenness	PageRank	Qualitative
0	-1.77	0.244	0.058	0.186	0.285	35724	5.9420E-05	Tight Aggressive
1	-4.57	0.370	0.208	0.162	0.284	19882	5.5610E-05	Loose Passive
2	-7.86	0.507	0.339	0.168	0.285	6376	3.6100E-05	Fish
3	-3.65	0.286	0.103	0.183	0.284	25165	5.1940E-05	Loose Aggressive
4	-4.49	0.241	0.074	0.167	0.283	19414	5.4160E-05	Tight Aggressive
Stats with Betweenness								
Cluster	Winrate	VPIP	Agg Delta	PFR	WTSD	Betweenness	PageRank	Qualitative
0	-4.25	0.243	0.076	0.167	0.283	17928	5.2050E-05	Tight Aggressive
1	-3.15	0.290	0.106	0.184	0.284	27046	5.3530E-05	Loose Aggressive
2	-4.57	0.373	0.213	0.160	0.284	20732	5.6890E-05	Loose Passive
3	-2.95	0.244	0.062	0.182	0.285	31840	5.8000E-05	Tight Aggressive
4	-7.86	0.508	0.340	0.168	0.284	6467	3.6250E-05	Fish

Table 1. Resulting Cluster Centers and Qualitative Style Assignment









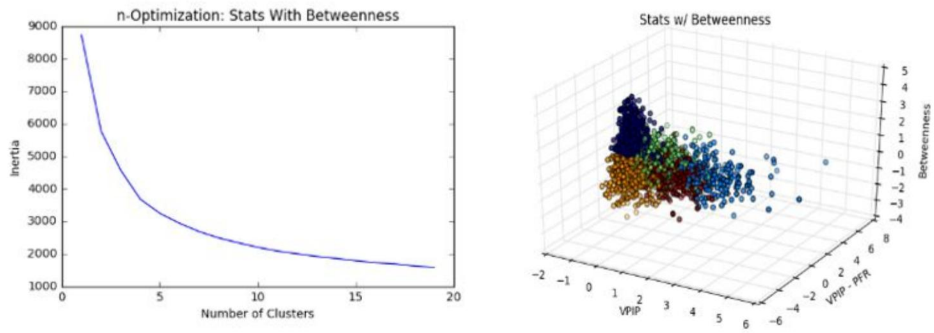


Figure 5. Inertia Curves and Cluster Assignments Using Three K-Means Models

The resulting qualitative analysis from the style summary intuition aligns quite well. There is some divergence from the playing style theory intuition, but considering that these hands all come from a faster-paced format (6 maximum players per table), these divergences should be expected - players will generally engage in more hands. Also, because link creation in the graph is probabilistic and per hand, players have more opportunities to add links: centrality is correlated to the total number of hands played. From this fact, the betweenness centrality, normalized by total number of hands played per cluster confirms player strategy intuitions.

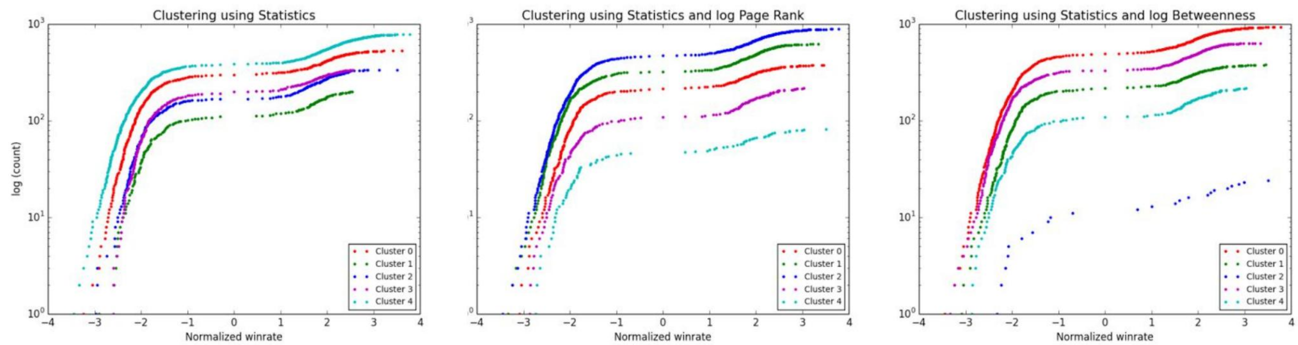


Figure 6: CDF of winrate (L to R: clustering with Statistics only, with PageRank and with Betweenness)

Comparing the Cumulative Distribution Function (CDF) of the winrate for clustering using the three methods we observe that the algorithm that incorporates the Pagerank and Betweenness has a better separation than the algorithms using only statistics. Note, only players with 1000 hands or more were considered in the clustering. In these graphs, the CDF's that are lower have less losses and are hence better. The clustering algorithm using betweenness is able to isolate a segment of players that had the least losses as a group (note that this may be because it is also a smaller set of players).

Stats Only		Stats with PageRank		Stats with Betweenness	
Betweenness/Hand	Qualitative	Betweenness/Hand	Qualitative	Betweenness/Hand	Qualitative
7.810	Tight Aggressive	3.216	Tight Aggressive	36.560	Tight Aggressive
40.996	Fish	83.452	Loose Passive	58.169	Loose Aggressive
5.751	Tight Aggressive	35.529	Fish	86.713	Loose Passive
80.633	Loose Passive	52.276	Loose Aggressive	3.615	Tight Aggressive
40.582	Loose Aggressive	32.251	Tight Aggressive	35.617	Fish

Table 2. Betweenness per Hand Played in Each K-Means Model

The most consistent losers in the graph have the highest linear approximation of betweenness rate of growth. The loose aggressive player is the only divergence from correlating betweenness/hand to winrate. This player appears to be balancing selectivity and aggressiveness, which will allow this player to be untrusted by others at the table and make it to showdown with strong hands. Because this player has a higher average VPIP, but a par WTSD, this player is getting to the showdown more often than less active players, but less than the loose passive and fish players. This cluster average representation of the loose aggressive player is probably a hybrid of the loose aggressive and (good) tight aggressive players discussed above.

Determining Predictive Significance of the Clusters

Stats - Only Clustering															
	p-value					Average Winrate (bb/100)					Total Hands Played				
src\dst	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
0	1.000	0.001	0.001	0.001	0.003	-1.01	0.91	-1.57	-0.48	-3.07	361224	7248	984729	20654	44057
1	0.001	0.472	0.001	0.022	0.013	0.00	0.47	1.58	-160.22	-21.66	0	147	23095	93	425
2	0.126	0.817	0.020	0.661	0.970	0.00	0.00	0.51	4.35	-0.86	0	0	2819051	65829	134731
3	0.003	0.506	0.001	0.206	0.001	0.00	0.00	0.00	-2.78	18.66	0	0	0	1004	555
4	0.007	0.674	0.015	0.464	0.243	0.00	0.00	0.00	0.00	1.34	0	0	0	0	5853
PageRank Clustering															
	p-value					Average Winrate (bb/100)					Total Hands Played				
src\dst	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
0	1.000	0.001	0.001	0.001	0.001	0.64	5.40	-1.92	5.12	3.09	4295389	81059	34753	229168	346600
1	0.001	0.031	0.253	0.036	0.001	0.00	-5.19	9.79	-8.44	-5.13	0	926	150	1178	1979
2	0.001	0.008	0.452	0.001	0.002	0.00	0.00	2.92	10.54	45.95	0	0	147	655	546
3	0.001	0.001	0.150	0.111	0.001	0.00	0.00	0.00	-4.52	-0.07	0	0	0	9324	12404
4	0.001	0.003	0.227	0.839	0.031	0.00	0.00	0.00	0.00	-2.73	0	0	0	0	27341
Betweenness Clustering															
	p-value					Average Winrate (bb/100)					Total Hands Played				
src\dst	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
0	1.000	0.006	0.401	0.001	0.296	-0.23	16.82	5.68	-0.88	-58.84	15392	7630	1490	259215	198
1	0.064	0.110	0.604	0.485	0.212	0.00	-4.79	22.49	-4.76	-9.30	0	6799	968	205401	520
2	0.554	0.070	0.680	0.001	0.840	0.00	0.00	-5.56	-5.61	9.79	0	0	926	78712	150
3	0.152	0.556	0.007	0.000	0.367	0.00	0.00	0.00	0.42	-2.23	0	0	0	4548043	35118
4	0.280	0.001	0.303	0.001	0.711	0.00	0.00	0.00	0.00	2.83	0	0	0	0	147

Table 3. Significance Test Results per Interaction per Model

Each of the clusters was partitioned against all opponent clusters and simulated to produce an encoded linear model to predict any significant change in the expected winrate when one cluster plays against a specific opponent style. The coefficients of the resulting cluster versus cluster models were analyzed for their standard error and p-value. Table 3 shows the results of the model when the top axis (destination node) engages with the index (source node).

The PageRank cluster models display the highest proportion of significant interactions. This result is expected, as PageRank in this context is a measure of having many in-degrees (winning edges) from those who also have high in-degrees and fewer out-degrees (losing edges). Table 4 shows the qualitative analysis of 'who wins from whom.'

Stats - Only Clustering			PageRank Clustering					Betweenness Clustering				
Qualitative Analysis			Qualitative Analysis					Qualitative Analysis				
Qualitative	Wins From	Neutral With	Qualitative	Count	Wins From	Neutral With	Loses To	Qualitative	Count	Wins From	Neutral With	Loses To
Tight Aggressive	TAG, LAG	LP, Fish	Tight Aggressive	449	TAG, LP, Fish		TAG, TP, LAG	Tight Aggressive	579	TAG, Fish	TAG	TP, LP, LAG
Fish	TP, Fish, LP, LAG	TAG	Loose Passive	358	TAG, LAG, TP		Fish, LP, TAG	Loose Aggressive	476	TP, TAG, Fish		LP, LAG
Tight Aggressive	Fish, TAG, LAG	TAG, LAG	Fish	202	LP, Fish		TAG, LAG, TP	Loose Passive	344	TAG, LAG, TP		LP, Fish
Loose Passive	TAG, LAG	TAG	Loose Aggressive	525	TP, TAG, Fish	TAG	LP, LAG	Tight Aggressive	582	TAG, Fish		TP, LP, LAG
Loose Aggressive	LP, LAG	LP	Tight Aggressive	646	TAG, Fish		LP, TP, LAG	Fish	199	LP, Fish		TP, TAG, LAG

Table 4. Qualitative Analysis of Significance and Winrates Between Clusters

In the tabled results, the player strategy intuition developed above is best described by PageRank and Betweenness cluster assignments. This output captures a selective mentality of players that the top-line statistics are not capable of. By seeing how 'thoughtful' these player types are about how they decide to connect to the graph through centrality measures, we are able to distinguish between player styles that are able to take advantage of other strategies and avoid those whose strategy disables their own.

Conclusion

In this paper, we explored the online poker dataset to uncover any differences that may exist between elite, moderate, and poor poker players based on how their game selection. We used clustering algorithms using key summary statistics along with their network statistics - betweenness and pagerank - to segment the users that allowed us to predict how money will flow on the network of players. We then analyzed the player styles and strategies which gave us insights to the significance of these clusters. Our results indicate that centrality measures of a player are a significant predictor of their exchanges with others, when combined with their top-line statistics. We used centrality measures to explain behaviors of players beyond what current poker statistics are capable of.

References

1. Little, J. Strategies for Beating Small Stakes Poker Cash Games. Jonathan Little. NY (2015).
2. Little, J. On Live No-Limit Cash Games. D&B Publishing. East Sussex, UK (2014).
3. Carlin Eng, Chetan Sharma, *Examining Online Poker as a Social Network,* CS 322: Social and Information Network Analysis Professor Jure Leskovec, Autumn 2009
4. Neil Shah, Alex Beutel, Bryan Hooi, Leman Akoglu, Stephan Gunnemann, Disha Makhija, Mohit Kumar, Christos Faloutsos, "EdgeCentric: Anomaly Detection in Edge-Attributed Networks" (2015)
5. Jure Leskovec, Daniel Huttenlocher, Jon Kleinberg. "Signed Networks in Social Media", Proceedings of the SIGCHI Conference on Human Factors in Computing Systems
6. Shi-Jen Lin , Yi-Ying Jheng, Cheng-Hsien Yu. "Combining ranking concept and social network analysis to detect collusive groups in online auctions"