Passing Power: Social Network Insights for Scoring Success in the NBA

Lucas Holliday, Evan Le, Josie Yan, and Michael Edgar

McCormick School of Engineering

Northwestern University

IEMS 341: Social Network Analysis

Dr. Noshir Contractor

June 11, 2025

Summary of Results

We used Social Network Analysis (SNA) techniques to examine how passing dynamics influence team success and how a player's network position and individual attributes affect their scoring performance across all 30 National Basketball Association (NBA) teams during the 2023–24 and 2024–25 seasons. Our goal was to generate actionable insights for Chicago Bulls Head Coach Billy Donovan and General Manager Marc Eversley to support coaching strategies, roster construction, and player recruitment and development.

Our project evolved from our initial proposal that focused on using Amazon book co-purchasing data to improve inventory curation and book sales for independent bookstores. After running exploratory analysis on sample data, we realized that the metadata included inconsistent genre labeling, making it unwieldy to capture behavioral buying trends solely with book titles. As we learned more about centrality, contagion, and the Autologistic Actor Attribute Model (ALAAM) in Lab 5, we decided to pivot to a sports analytics project that would allow us to apply these metrics. Furthermore, with the NBA playoffs ongoing, analyzing player dynamics and recommending data-driven coaching strategies felt more relevant and interesting to us.

We retrieved the NBA dataset for the 2023-24 and 2024-25 seasons using the nba_api GitHub package, an API Client for www.nba.com. Through a Python script, we collected directed and weighted passing and assist interactions among the top ten players per team, selected based on total minutes played across each season. This approach generated a total of 60 networks (30 teams for each of the two seasons), allowing for comparisons of passing and assists structures within and across franchises.

For descriptive statistics, each network consisted of a complete graph of 10 nodes since every player made at least one pass to each teammate. For example, edges in the passing network for the Chicago Bulls in the 2024-25 season had on average weight 163.3 with a high of 1,064 passes from Josh Giddey to Nikola Vučević. Later, when collecting attribute data for the ALAAM, we used the same API and a Python script to retrieve player characteristics such as average points per game, age, team tenure, position, and draft pick.

Visual graphs of the passing networks show how The Denver Nuggets' offense displays a hub-and-spoke pattern centered on MVP Nikola Jokić, reflecting a star-driven system with high centrality. In contrast, the Portland Trail Blazers, a team rebuilding after trading their star Damian Lilliard, exhibit a more decentralized network with passing more evenly distributed across players (see Figure 1).

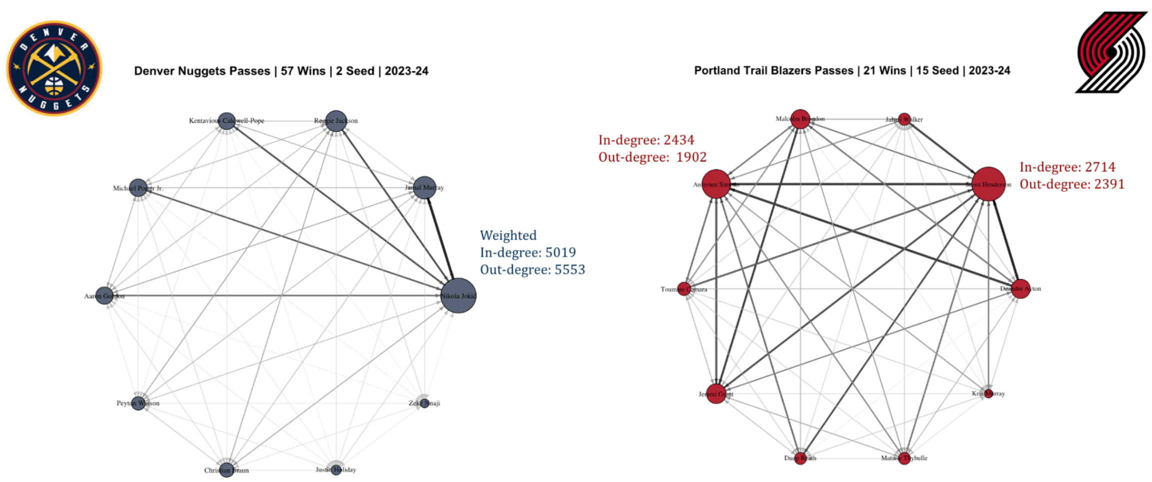

Figure 1. Differences in Passing Network Structures (Denver Nuggets vs Portland Blazers)

One strength of the dataset is that it curates data from the NBA, ensuring a high level of reliability and accuracy. It also captures detailed on-ball passing among the league's most active players, offering insights into how teams move the ball and create scoring opportunities during games. However, the dataset does not account for off-the-ball movement or defensive strategies, and our approach does not include game factors such as score differential, game pace, or time remaining. Additionally, it excludes the other five players on a NBA team roster and fails to account for variability introduced by line-up changes and substitutions.

Since our class analysis techniques only applied to unweighted graphs, we binarized the network by retaining only edges with weights above the average (see Figure 2). While this simplification allowed us to apply methods covered in class without requiring extensive additional research, it came at a cost. Binarization removed the actual number of passes between players, collapsing an interaction of hundreds of passes over a season into a simple "yes/no" link. As a result, a strong connection between two stars running a two-man game offense was treated identically to one just barely above the average, while all sub-average passing relationships were excluded entirely. This limitation reduced the level of detail of our analysis and prevented us from exploring how variations in pass volume and its distribution across players might influence overall team performance. In these binarized networks, each team had 10 nodes (representing the top 10 players) with an average of 30.6 directed edges per network. Across all teams' passing networks, the number of edges ranged from a low of 20 to a high of 41, reflecting variation in how concentrated or evenly distributed passing patterns were among different teams. The average network density was 0.34 and the average reciprocity was 0.52.

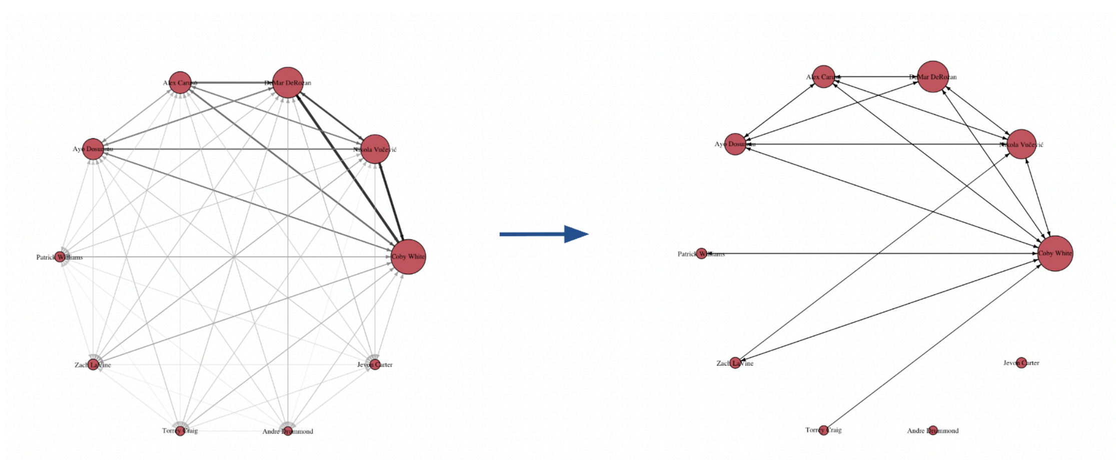

Figure 2. Chicago Bulls (2023-24) Passing Network Before and After Binarization

Questions

The first question we attempted to answer was what network properties are associated with high team performance. Curious about what differentiates great teams, we wanted to determine whether successful NBA teams share specific structural characteristics in their passing networks, such as higher reciprocity, greater density, or more evenly distributed interactions among players.

The second question we attempted to answer was what individual and relational metrics best predict high scoring performance. We wanted to understand whether there are underlying factors, such as age, team tenure, draft position or network position (e.g., how many teammates a player passes or receives passes from) that are more strongly correlated with being a high scorer.

Since the Chicago Bulls fell short of the playoffs after losing in the Play-In Tournament, these questions are important for the organization. Understanding how passing dynamics and player connectivity contribute to team performance can help optimize offensive strategies, lineups, and in-game decisions. Additionally, identifying the attributes that predict high scorers can improve scouting, roster construction, and player development, helping to build a more competitive team and a winning dynasty in the league.

Analysis

We conducted a network-level regression and ALAAM analysis to explore team-level passing structures and individual player attributes within a social network framework. We chose these methods because they allowed us to effectively address our research questions.

<u>Network-level Regression Analysis</u>

In the network-level regression, we examined how global properties such as edge count, density, average degree, average path length, centralization, and assortativity from team passing and team assist networks correlated with total wins using cor.test() in R. The analysis was performed using data from all NBA teams across the past two seasons (2023–24 and 2024–25).

We chose this analysis to explore the statistical relationships between network structure metrics and team success.

<u>ALAAM Analysis</u>

Our ALAAM analysis focused on modeling individual-level scoring outcomes, incorporating network and player specific attributes. As shown in Figure 3, we constructed a league-wide passing network composed of the top ten players per team of all 30 teams.

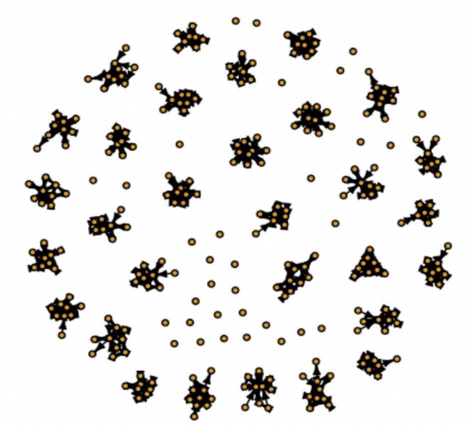

Figure 3. NBA Passing Network (2023-24)

Combining all teams into a single network increased the number of nodes and edges, which helped achieve the higher Effective Sample Size (ESS) scores needed for our analysis. Along with our adjacency matrix, our attribute data of each player included individual characteristics, such as age, team tenure, and draft position. Since there are 60 picks in the NBA draft, an undrafted player was assigned as pick 61. We created a binary variable to identify high scorers, defined as players averaging more than 20 points per game (PPG). The model included structural variables such as in-degree (passes received), out-degree (passes made), and reciprocity. We chose the ALAAM because it allows for the isolation of each variable's effect while accounting for the interdependencies that arise in network data.

We tested the following six hypotheses:

1. Players are less likely to be high scorers if they are closely connected to other high scorers (negative contagion effect).
2. Older players are more likely to be high scorers.
3. Players with longer team tenure are more likely to be high scorers.
4. Players drafted earlier (lower draft pick) are more likely to be high scorers.
5. Players who *pass* the ball to more teammates (higher out-degree) are more likely to be high scorers.
6. Players who *receive* the ball from more teammates (higher in-degree) are more likely to be high scorers.

Since all Effective Sample Size (ESS) values exceeded 200 (see Figure 4), the results demonstrate sufficient convergence and suggest that the simulation estimates are reliable. The iteration count of 40,000 was selected due to the large number of parameters involved. Additionally, since the ESS values did not exceed 10,000, the computation was not considered excessive or wasteful.

```
you have done  40000  iterations out of  40000
theta: -10.256 -1.499 0.242 0.475 0.421 -1.228 -2.616 -0.059 -4.086 0.991 4.065 0.121
summaries of the posterior draws:
                             mean          sd          ESS      SACF 10      SACF 30
intercept            -10.42155678  7.06905748 377.77657666   0.81435423   0.55603966
contagion             -1.50596747  0.77597994 414.75080728   0.79715241   0.53430935
age                    0.18471364  0.17043303 423.47124719   0.78823030   0.49470771
tenure_years           0.49030949  0.33293813 399.70742053   0.80397339   0.53799461
guard                 -2.73188886  1.96964808 342.36696296   0.82845406   0.56933987
prime_age              1.48010541  1.57299096 421.76473404   0.79154356   0.51388661
first_round           -1.07828417  3.80310083 286.01631723   0.85140366   0.62171738
draft_pick            -0.12468188  0.09763363 375.63616065   0.80747243   0.53222002
out.degree            -3.68504986  2.50078804 268.75547752   0.86612589   0.65989377
in.degree              2.28186567  0.98859751 407.55172553   0.80188634   0.54242886
reciprocity            2.73172101  2.36372490 332.65229170   0.83456831   0.59843948
transitive.triangles   0.13382643  0.28194685 258.74564357   0.86613406   0.65033567
```

Figure 4. Effective sample sizes (ESS) Model Results

In our Goodness-of-Fit test (see Figure 5), the p-values ranged from 0.078 (instar) to 0.253 (transitive cont.). Since these p-values were all well above our threshold of 0.05, we concluded that our model was a proper fit for the data.

```
Simulating GOF took  7.348319
Calculating statistics took  0.871665
                        obs     mean    p-val
intercept            50.000   48.654    0.153
simple cont.         44.000   41.028    0.165
recip cont.          22.000   20.148    0.151
indirect cont.      147.000  146.780    0.199
closedind cont.     143.000  131.778    0.164
transitive cont.     12.000   17.092    0.253
outdegree           243.000  230.548    0.110
indegree            293.000  274.628    0.087
reciprochation      238.000  226.274    0.114
instar              798.000  734.014    0.078
outstar             542.000  508.098    0.115
twopath            1306.000 1212.730    0.097
in3star            1296.000 1190.112    0.099
out3star            707.000  664.570    0.143
transitive          566.000  542.724    0.147
cyclic              617.000  584.874    0.111
indirect            731.000  708.300    0.158
excl.indirect        95.000  101.318    0.170
```

Figure 5. Goodness-of-Fit Test Results

Findings

Network-level Regression Analysis

Regression analysis revealed several relationships between network structure and team success. In passing networks, among all global network properties analyzed, in-degree assortativity had the lowest p-value (0.16), though it was not statistically significant at the 0.05 level. The estimate of -0.18 suggests negative in-degree assortativity, meaning that on more successful teams, players who receive a lot of passes tend to receive them from teammates who receive fewer passes. Passes are often funneled toward a few central receivers, typically the stars on each team. Interestingly, when isolating data from only Western Conference teams over the past two seasons, we found statistically significant correlations between total team wins and three network metrics: (1) edge count, (2) density, and (3) average degree. Since each passing network contains the same number of nodes (ten players), these metrics are mathematically related and produced nearly identical p-values of 0.018 and correlation coefficients of -0.42. This negative relationship suggests that less densely connected teams in the Western Conference tended to win more games, which aligns with the dominance of high usage stars such as LeBron James (4x MVP), Steph Curry (2x MVP), Nikola Jokic (3x MVP), Shai Gilgous Alexander (1x MVP), James Harden (1x MVP), Kevin Durant (1x MVP), and Luka Doncic. Many playoff teams in the Western Conference appear to be built around one or two focal players. In contrast, the same analysis in the Eastern Conference revealed a positive correlation of 0.23 between density and wins, suggesting a more balanced pattern of ball movement. However, this relationship was not statistically significant with a p-value of 0.22. Compared to the west, the Eastern Conference only has two current players, Giannis (2x MVP) and Embiid (1x MVP), who have won MVPs.

The assist networks produced a more robust result. Across all NBA teams, those with shorter average path lengths in their assist networks (see Figure 6) tended to win more games.

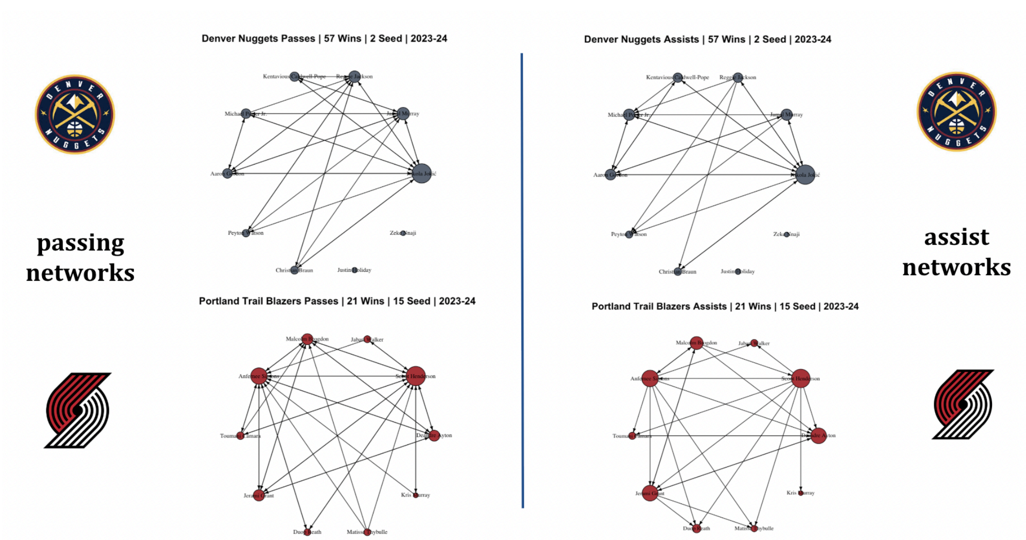

Figure 6. Examples of Passing and Assist Networks

This relationship was statistically significant (p-value = 0.0099) with a Pearson correlation coefficient of –0.33, indicating a moderate negative association. In other words, teams that move the ball efficiently and involve more players in their offensive sets (requiring fewer steps to reach a scorer) are more successful. A more decentralized assist structure makes it more challenging for defenses to collapse on a single playmaker since the offense flows quickly and unpredictably through multiple contributors.

ALAAM Analysis

ALAAM analysis uncovered several interesting insights into the dynamics of scoring within NBA teams. We identified a statistically significant contagion effect for high scoring. The 95% credibility interval for this effect was [-2.91, -0.157], which excludes zero, indicating a conclusive result. The negative estimate (-1.378) and odds ratio of 0.252 suggest that a player is approximately 75 percent *less* likely to be a high scorer if they are closely connected (by passing ties) to other high scorers. This finding challenges the common assumption that proximity to dominant scorers increases individual output. Support players may defer to the star players which reduces their own offensive play, possibly due to fewer touches, lower shot attempts, or psychological deference to higher-status teammates.

In contrast, we found a statistically significant strong positive relationship between inbound passing connections and scoring performance. The effect of in-degree, defined as the number of teammates passing to a player, had a 95% credibility interval of [0.647, 4.499], with a positive estimate of 2.271. The corresponding odds ratio of 9.689 indicates that players who receive passes from more teammates are nearly 870% more likely to be high scorers. This result highlights the importance of inbound ball movement in driving individual scoring success.

Other variables were not statistically conclusive, but showed directional trends. Team tenure had a positive estimate (0.466) with a 95% credibility interval of [-0.097, 1.10], suggesting that players who have been with the team longer may be more likely to score, although we fail to reject the null hypothesis. Similarly, draft position had a negative estimate (-0.125), indicating that players picked earlier in the draft may be more likely to score more, but the credibility interval [-0.346, 0.044] includes zero, making this effect inconclusive. Lastly, out-degree (passes made to others) had a negative estimate (-3.302) with a wide 95% credibility interval of [-8.516, 0.231], indicating a high uncertainty and no clear relationship between outbound passing and scoring.

Implications and Recommendations

The statistically significant findings regarding negative contagion and positive in-degree provide actionable insights into how network structure can support individual offensive output. The negative contagion effect reveals that clustering high scorers in the same lineup may unintentionally suppress overall scoring since players may defer to established scorers rather than assert their own offensive roles. To counteract this dynamic, we recommend that the Chicago Bulls stagger their top scorers across rotations to create a more balanced offensive presence throughout the game and to encourage supporting players to maintain an aggressive scoring mindset.

The finding that lower average path length in assist networks correlates with team wins and the strong positive effect of inbound passing (in-degree) on scoring highlights the value of ball movement from multiple teammates in creating diverse scoring opportunities since denser passing networks may enable more efficient ball movement and decision-making on the court. Furthermore, players who receive passes from a wider range of teammates are significantly more likely to be high scorers, highlighting the importance of unselfish play and offensive strategies that increase connectivity. As a result, we recommend that the Chicago Bulls create offensive plays that move the ball through multiple players with a focus on high-percentage passes, rather than relying on individual scorers.

While the results for team tenure and draft position were not statistically conclusive, the directional trends suggest that longer team tenure and player development may positively influence scoring productivity. Based on these insights, we recommend that the Chicago Bulls prioritize long-term development plans, especially for players selected outside of the first round or lottery. Focusing on internal growth and player continuity may result in higher team performance than pursuing external star power alone.

Reflection

The regression and ALAAM network modeling allowed us to address the questions we set out to explore, specifically which social network and individual attributes most strongly predict high scoring success among NBA players and how these insights can inform team strategic decisions for the Chicago Bulls. Not surprisingly, players with higher in-degree centrality (those who receive passes more from teammates) were significantly more likely to be high scorers. This result confirmed our initial hypothesis that being well-integrated into the team's passing structure facilitates more scoring opportunities, reinforcing the value of ball movement and inclusive offensive plays.

However, we were surprised by the statistically significant negative contagion effect. Contrary to our expectations, players who were closely connected to high scorers were less likely to be high scorers themselves. This result ran counter to our initial assumption that playing alongside star players would improve individual performance. Proximity to dominant scorers likely leads other players in the network to take on more passive offensive roles, limiting scoring opportunities. We were also struck by the directional, yet inconclusive, trends in team tenure and draft position. The possibility that longer team tenure might be more strongly correlated with scoring success than draft status raises questions about the value of player development and team culture. Future research using larger samples, additional seasons, or more detailed in-game data would allow us to better understand how these network structures specifically shape individual outcomes and team performance.

One limitation of our analysis was the need to binarize the passing and assist networks due to the methodological constraints of the models covered in class. This simplification excluded valuable information about the volume and strength of interactions. Future research could build on our work by leveraging the original weighted data, applying models such as weighted ERGMs to better capture patterns and interaction strengths. Another opportunity lies in integrating the passing and assist networks. While we analyzed these networks separately, we did not explore how they interact, for example, by identifying which passes most frequently lead to assists. Analyzing pass-to-assist conversion rates could help differentiate general ball movers from highly effective facilitators. Additionally, incorporating Relational Event Models (REMs) with timestamped in-game events could help capture the temporal and sequential aspects of passing behavior. Together, these directions would offer a more comprehensive understanding of how passing patterns influence individual scoring outcomes and overall team success.

References

Borgatti, S.P., Everett, M. G., Johnson, J. C., & Agneessens, F. (2022) *Analyzing social networks using R*. SAGE Publications.

Contractor, N. & Forbush, E. (2017). Networks. In C.R. Scott, J.R. Barker, T. Kuhn, J. Keyton, P.K. Turner, & L.K. Lewis (Eds.), *The international encyclopedia of organizational communication*, Wiley.

Prell, C. (2012). *Social network analysis: History, theory and methodology*. SAGE Publications.