# FBS Networks in the College Football Playoff Era

**Eli Gnesin**

Department of Data Science, College of William and Mary, Williamsburg, VA

## Background or Abstract

In 2014, the NCAA Division I Football Bowl Subdivision (FBS) began a new playoff format, the College Football Playoff (CFP). In the eight years since, five different teams have won, but only thirteen different teams have played in the Playoff, leading to discussions of expansion[1,2,3]. This project examines the FBS during this era, at both the conference and team level, to uncover which team(s) have been most central to the FBS in this era.

## Research Questions

1) Which team(s) are most important or influential to the FBS in the CFP era?
2) To what extent are playoff teams more central to the CFP era?

## Related Research

FiveThirtyEight did an analysis of CFP field sizes by considering an optimized way of finding the "best" team in a season, which determined that 12 teams would be enough in most years[2]. In February 2022, however, plans to expand the CFP to 12 teams fell apart with pushback from three conferences[3].

## Data

The data was collected from Sports-Reference's College Football database and includes the FBS schedule and standings for eight seasons between 2014 and 2021[4]. Data was collected by copying the csv data from the standings and schedule page for each season into csv Files. Overall, there are 1034 team seasons and 6715 games, for an average of 129 teams and 840 games per season.

| Season | # of Teams | # of Games |
|---|---|---|
| 2014 | 128 | 868 |
| 2015 | 128 | 871 |
| 2016 | 128 | 873 |
| 2017 | 130 | 874 |
| 2018 | 130 | 884 |
| 2019 | 130 | 888 |
| 2020 | 130 | 570 |
| 2021 | 130 | 887 |
| Total | 1034 | 6715 |

**Table 1**: Metadata from the seasons in the dataset.

## Preprocessing

Five steps were taken to preprocess the data before analysis or visualization:

1. Standardize team names across schedule/standings.
2. Remove AP rankings from team names in schedule.
3. Add conference affiliation to each team in schedule.
4. Add year for each game in schedule/team in standings.
5. Concatenate single dataframe for standings/schedule.

Later steps taken to create conference-based network visualizations:

1. Create empty adjacency matrix with conference rows and columns.
2. Iterate over schedule by game, increasing cell at index (winner, loser).
3. Reset diagonal to zero to ignore intra-conference games.

**Table 2**: The non-conference games adjacency matrix for the entire dataset, with winners by row and losers by column.

| Winner | Sun Belt | MAC | Pac-12 | Big Ten | SEC | MWC | American | Ind | ACC | CUSA | FCS | Big 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Sun Belt | 0 | 21 | 0 | 1 | 4 | 12 | 5 | 16 | 2 | 29 | 65 | 5 |
| MAC | 15 | 0 | 1 | 14 | 1 | 16 | 4 | 10 | 3 | 16 | 71 | 5 |
| Pac-12 | 9 | 7 | 0 | 23 | 5 | 53 | 3 | 18 | 11 | 9 | 54 | 10 |
| Big Ten | 12 | 59 | 25 | 0 | 14 | 26 | 22 | 18 | 29 | 23 | 41 | 14 |
| SEC | 52 | 25 | 7 | 21 | 0 | 16 | 22 | 18 | 45 | 51 | 94 | 21 |
| MWC | 19 | 15 | 27 | 1 | 4 | 0 | 14 | 36 | 5 | 16 | 71 | 0 |
| American | 19 | 15 | 6 | 8 | 9 | 11 | 0 | 21 | 21 | 31 | 69 | 7 |
| Ind | 16 | 19 | 18 | 10 | 6 | 29 | 27 | 0 | 34 | 27 | 37 | 3 |
| ACC | 20 | 27 | 2 | 19 | 35 | 4 | 33 | 25 | 0 | 32 | 103 | 11 |
| CUSA | 27 | 28 | 0 | 3 | 5 | 8 | 18 | 21 | 5 | 0 | 83 | 1 |
| FCS | 8 | 8 | 4 | 2 | 3 | 7 | 6 | 7 | 3 | 8 | 0 | 5 |
| Big 12 | 7 | 9 | 18 | 9 | 18 | 7 | 23 | 6 | 8 | 29 | 55 | 0 |

## Methods or Analysis

To analyze and explore this dataset, I relied on network analysis using the NetworkX package in Python, with a focus on centrality measures for various data subsets.

$$C(u) = \frac{n-1}{\sum_{v=1}^{n-1} d(v,u)}$$

$$c_B(v) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}$$

**Figure 3**: Formulas for Closeness and Betweenness centrality in NetworkX

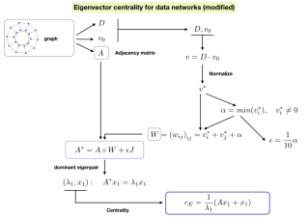


**Figure 4**: Eigenvector centrality[5]

## Results

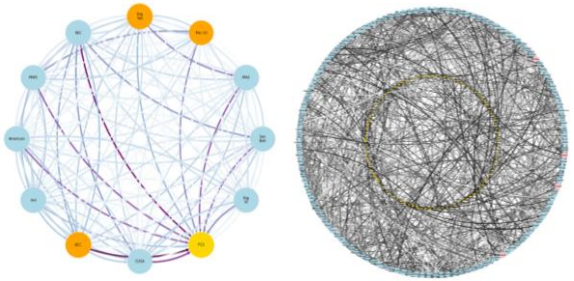

**Figure 5:** (a) The overall non-conference network. Orange indicates votes against playoff expansion, gold indicates FCS (no vote), (b) The schedule network for 2021. Playoff teams are shaded pink and FCS teams are shaded gold.

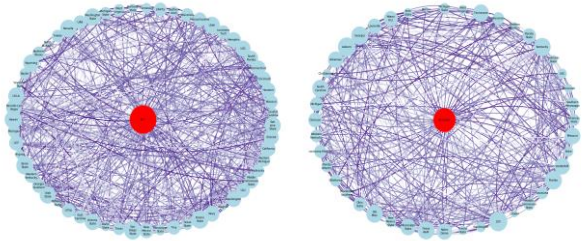

**Figure 6:** (a) BYU ego network (top degree/closeness centrality), (b) Alabama ego network(top eigenvector centrality)

| | Degree | Betweenness | Closeness | Eigenvector |
|---|---|---|---|---|
| **Playoff** | 0.199359 | 0.005079 | 0.481904 | 0.139301 |
| **Non-Playoff** | 0.111440 | 0.004324 | 0.289716 | 0.038558 |

**Table 7**: Average centrality scores for 13 schools that have made the playoff vs. all schools that have not.

| | Degree | Betweenness | Closeness | Eigenvector |
|---|---|---|---|---|
| **Team #1** | BYU | Army | BYU | Alabama |
| **Team #2** | Army | BYU | Alabama | Ohio State |
| **Team #3** | Massachusetts | Massachusetts | LSU | Notre Dame |

**Table 8**: The top 3 schools by each main centrality measure

## Discussion

Understanding which team is "most influential" is metric dependent. The main FBS independents (BYU, Army, Massachusetts, Notre Dame) are central to the FBS in the era by playing more different teams over time. Conversely, Alabama has high rank by "influence" centrality metrics such as eigenvector centrality. Further, the average score across each centrality metric is higher for the 13 playoff teams than the non-playoff teams, suggesting that they are, in fact, more "central" to the CFP era. However, these metrics are predicated on ignoring the "parallel edges" from multiple games between two teams, a major source of bias.

## Conclusion

Though there is much analysis of the FBS during the CFP era from various perspectives, little to none of it takes a network analysis perspective to the topic. This perspective allows for new analysis of centrality and communities in the FBS, for the purposes of understanding influence and importance for teams.

## Future Work

I. Consider future realignment beginning with 2022-2023 season.
II. Community detection analysis at the team and conference level in expansion context.
III. Expand analysis to include Football Championship Subdivision (FCS) teams.
IV. Measure centrality metrics when including parallel edges.

## References

1. Sidearm Sports 2019 *College Football Playoff*
2. Wissner-Gross 2021 *FiveThirtyEight*
3. Auerbach et al. 2022 *The Athletic*
4. Sports-Reference n.d. *College Football Reference*
5. Pedroche et al. 2019 *Symmetry*