The Madness of March Madness

An OR Approach to Team Rankings in NCAA

Tiana Marrese, Mary St. John, Tyler Chapin, Allison Klenk

Operations Research II

12/13/17

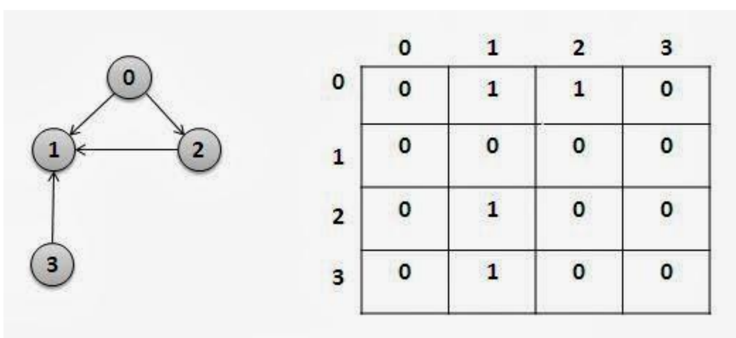
Introduction/Motivation

Each spring, the United States goes crazy over college basketball. The NCAA Division I Men's Basketball Tournament, known as March Madness, sees to determine the best of the best of college basketball through three weeks of single-elimination games. Out of 134 total teams, 68 are chosen to participate in March Madness through two steps. The first 32 teams are automatically a part of the tournament after winning their respective conference and the remaining 36 are chosen by a selection committee based on their performance that year. The 64 teams are then split up into four regions and are ranked from 1-16. Typically, higher-ranked teams begin playing lower-ranked teams until upsets occur which make the tournament unpredictable. Each year, millions of people create their own bracket in an attempt to predict the result of each game. It should be noted that there is a 1 in 9.2 quintillion chance of creating the perfect bracket.

The selection committee currently uses the Rating Percentage Index (RPI) to aid the selection of the final 36 teams. RPI takes into account each team's winning percentage, their opponents' winning percentages, and their opponents' opponents' winning percentages. Although this gives an idea of how "good" a team is, it also emphasizes the importance of a good schedule. Even with RPI, the committee relies a great deal on human reasoning, which is convoluted with biases and self-interest. One major source of this bias lies in the idea that the NCAA makes roughly 85% of their revenue from March Madness, so it is to the committee's benefit to select interesting and exciting matchups that provide entertainment and thus generate revenue. In our project, we sought to replicate the seeding of the teams by creating a new rating system that rid the process of unfair judgements. We began by questioning whether there was a way to represent teams and even out the playing field by generating a better March Madness bracket. This simple question led us to wonder whether we could extrapolate this to other sports teams and solve a major issue of college sport rankings in the United States.

Concepts

The PageRank algorithm was invented in 1996 by Larry Page and Sergey Brin, who would then go on to co-found Google. The goal was simple -- they wanted to be able to procedurally compute a well-defined ranking for a huge, incredibly interlaced span of web pages, but inferring value from page contents was a very difficult problem, both analytically speaking and in terms of result quality. The idea behind PageRank was to circumvent those difficulties by examining the *connections* between pages and glean the information from that -- which brings us to our project. Ranking basketball teams suffers from many of the same problems when trying to judge team-specific qualities; there's a lot of complex and frankly non-objective analysis that comes from examining a sport directly. By analyzing the network of head-to-head victories between teams, we believe we can gleam better information.



Adjacency Matrix Representation of Directed Graph

In both cases, the actual algorithm is deceptively simple.

The first step is to reduce the network to a single directed graph (which we'll represent with an adjacency matrix we'll call **A**). In our case this boiled down to the head to head record of team *i* vs team *j* being placed in $\mathbf{A}_{ij}$ of the adjacency matrix (note that, naturally, all the diagonal entries would be 0). This stage is where any weighting methods we're using to individual matches would be applied -- if for example instead of just counting wins we wanted to factor in the score of those wins, the weighting would be applied to that edge in the matrix. Next, you define a base vector **v** -- typically for some *n* teams, it would just be $(1/n)\mathbf{1}^n$, that is to say the n-length unit vector scaled down to sum to 1. However, if you *did* want to weight to prefer individual teams, this is where that would happen.

$$\mathbf{v} = \begin{pmatrix} 0.25 \\ 0.25 \\ 0.25 \\ 0.25 \end{pmatrix}, \quad \mathbf{Av} = \begin{pmatrix} 0.37 \\ 0.08 \\ 0.33 \\ 0.20 \end{pmatrix}, \quad \mathbf{A}^2\mathbf{v} = \mathbf{A}(\mathbf{Av}) = \mathbf{A}\begin{pmatrix} 0.37 \\ 0.08 \\ 0.33 \\ 0.20 \end{pmatrix} = \begin{pmatrix} 0.43 \\ 0.12 \\ 0.27 \\ 0.16 \end{pmatrix}$$

$$\mathbf{A}^3\mathbf{v} = \begin{pmatrix} 0.35 \\ 0.14 \\ 0.29 \\ 0.20 \end{pmatrix}, \quad \mathbf{A}^4\mathbf{v} = \begin{pmatrix} 0.39 \\ 0.11 \\ 0.29 \\ 0.19 \end{pmatrix}, \quad \mathbf{A}^5\mathbf{v} = \begin{pmatrix} 0.39 \\ 0.13 \\ 0.28 \\ 0.19 \end{pmatrix}$$

$$\mathbf{A}^6\mathbf{v} = \begin{pmatrix} 0.38 \\ 0.13 \\ 0.29 \\ 0.19 \end{pmatrix}, \quad \mathbf{A}^7\mathbf{v} = \begin{pmatrix} 0.38 \\ 0.12 \\ 0.29 \\ 0.19 \end{pmatrix}, \quad \mathbf{A}^8\mathbf{v} = \begin{pmatrix} 0.38 \\ 0.12 \\ 0.29 \\ 0.19 \end{pmatrix}$$

Finally, we reach the procedural part. Observe this example process:

The first result is just **A** x **v**, the next result is just **A** left-multiplied by that, etc. You can continue this process indefinitely, and can attain convergence to within arbitrarily small bounds for the values. The end result gives a ranking for every team in the network, based on the size of the value in the corresponding row of the resulting vector.

It's an incredibly elegant process, but what's going on here? Conceptually, all the teams start on equal footing (with a normal base vector). With every iteration of the process, more "respect" is given to teams that win a lot as their weighting increases, but as those values get farther and farther from the baseline more and more weight is given for beating a team with real credibility, whereas picking up easy games against comparative no-names fails to impress -- *both of which are based in self-evident judgements*. This is, in fact, the kind of analysis one hopes to get from a committee of sports analysts, to help sift through the dearth of pre-bracket data with many incongruities in schedule difficulties -- and what's more, it's done objectively and without delving into obtuse and inscrutable analysis with so many moving parts that any verification of statistical relevance would be hopeless. Indeed, after an examination of the process it seems an extremely natural way to approach the problem.

Results

For this project we considered four separate variations of the PageRank algorithm to generate a numerated list of the teams. It is recognized that there are an extremely large amount of variations within college rankings right now due to factors including popular vote, conference setup, and human bias. Our list aims to rank the teams on a completely fair scale, disregarding all of these external factors. We believe the best data to rank teams should be based purely on scoring. Scores of games provide a factual, untainted indication of which team was better in the match. It is with this line of reasoning that our base case is a PageRank algorithm with an evenly distributed weight vector across teams. To achieve this we created a large matrix of all the teams then added a 1 in the space in which the column team beat the row team. These instances were then weighted by how many times the losing team lost in the season. Then using the equally distributed importance vector we computed the rank. This logic follows closely with a basic PageRank set up, for more insight refer to "Concepts" portion of this write up.

To compare with our base case, we explored three different weighting techniques in the adjacency matrix. The first approach is very similar to the second approach, except that instead of having a one where the column team beat the row team, we have the number of times the column team beat the row team. The second weighting method is the difference in points between the winning column team and the losing column team of their most recent head-to-head matching. The last weighting is similar to the second ranking, but instead of a difference in points, we report the ratio of points won by the column team to the points won by the losing row team. Taking into account the ratio allows a clearer understanding of which teams were stronger on the court and also controls for high scoring and low scoring games which the point difference metric does not.

These methods produced results similar to the tournament seedings. Most of the top few seeds from all regions appear in the top 40 ranking; however, there are some schools, like Gonzaga that we rank far below where they should be. They were a number one seed going into the tournament and finished second overall, and our ranking puts them at best 19th (number of wins) and at worst 33rd (score difference). Below is a table of our results. The first column denotes the seed that the team in the second column entered the tournament with. There are four teams ranked with each number because there is one team from each region with each ranking. The last four columns are the results of the PageRank algorithm with four different weighting strategies in the initial adjacency matrix. The top twelve teams are colored to easily track how our results compare to the official ranking. Most of the top twelve teams are seated near where they would be in the final bracket. Because the metrics produce very similar results, we propose that the standard ranking is sufficient to keep the model simple.

Table 1

	Tournament Seeds		Standard	Number of Wins	Score Difference	Score Ratio
1	Villanova		Butler	Butler	North Carolina	Butler
1	Gonzaga		North Carolina	Villanova	West Virginia	North Carolina

1	Kansas		Duke	Duke	Butler	Duke
1	North Carolina		Villanova	Creighton	Virginia	West Virginia
2	Duke		Kansas	Kansas	Duke	Villanova
2	Arizona		Florida St	Marquette	Baylor	Kansas
2	Louisville		Marquette	North Carolina	Villanova	Virginia
2	Kentucky		West Virginia	West Virginia	Louisville	Baylor
3	Baylor		Louisville	Iowa St	Florida St	Louisville
3	Florida St		Baylor	Florida St	Miami FL	Florida St
3	Oregon		Iowa St	Michigan	Creighton	Marquette
3	UCLA		Notre Dame	BYU	Iowa St	Iowa St
4	Florida		Virginia	Seton Hall	Michigan	Notre Dame
4	West Virginia		Seton Hall	Louisville	Georgia Tech	Wisconsin
4	Purdue		Wisconsin	Baylor	Wisconsin	Michigan
4	Butler		UCLA	Providence	Georgetown	Seton Hall
5	Virginia		Minnesota	Notre Dame	Oklahoma St	Creighton
5	Notre Dame		Providence	Virginia	Florida	Providence
5	Iowa St		Creighton	Gonzaga	Kansas	Miami FL
5	Minnesota		Michigan	Wisconsin	Notre Dame	BYU
6	SMU		BYU	Arizona	Marquette	Minnesota
6	Maryland		Oregon	UCLA	BYU	UCLA
6	Creighton		Miami FL	Xavier	Kentucky	Georgetown
6	Cincinnati		Xavier	Georgetown	Oregon	Oregon
7	South Carolina		Indiana	Oregon	Syracuse	Indiana
7	St Mary's CA		Georgetown	Minnesota	Providence	Xavier
7	Michigan		Arizona	Purdue	Kansas St	Arizona
7	Dayton		Maryland	Indiana	Indiana	Purdue
8	Wisconsin		Purdue	Maryland	Purdue	Gonzaga
8	Northwestern		Gonzaga	St John's	Xavier	Georgia Tech
8	Miami FL		Iowa	Miami FL	Iowa	Iowa

8	Arkansas		Syracuse	Michigan St	Seton Hall	Syracuse
9	Virginia Tech		Virginia Tech	Virginia Tech	Gonzaga	Maryland
9	Vanderbilt		Georgia Tech	Iowa	Virginia Tech	Kentucky
9	Michigan St		St John's	Syracuse	Minnesota	Virginia Tech
9	Seton Hall		Michigan St	Kentucky	Pittsburgh	Michigan St
10	Marquette		Kentucky	Kansas St	Arizona	St John's
10	VA Commonwealth		Pittsburgh	Northwestern	UCLA	Kansas St
10	Oklahoma St		Northwestern	TCU	Michigan St	Northwestern
10	Wichita St		USC	Illinois	St John's	Pittsburgh
11	Providence		Kansas St	Georgia Tech	Wake Forest	Florida
11	Xavier		TCU	Pittsburgh	Vanderbilt	TCU
11	Rhode Island		Florida	USC	Northwestern	USC
11	Kansas St		Ohio St	Florida	Oklahoma	Oklahoma St
12	USC		Texas Tech	Oklahoma St	South Carolina	Texas Tech
12	Princeton		Oklahoma St	Ohio St	TCU	Ohio St
12	Nevada		Wake Forest	Cincinnati	Illinois	Illinois
12	Wake Forest		Illinois	SMU	USC	Wake Forest
13	UNC Wilmington		Penn St	Vanderbilt	SMU	Cincinnati
13	Bucknell		Cincinnati	Penn St	Texas Tech	Penn St
13	Vermont		Nebraska	St Mary's CA	Maryland	Vanderbilt
13	MTSU		Vanderbilt	Wake Forest	Tennessee	Nebraska
14	ETSU		SMU	Texas Tech	St Mary's CA	SMU
14	FL Gulf Coast		South Carolina	Nebraska	Temple	South Carolina
14	Iona		NC State	South Carolina	Ohio St	Oklahoma
14	Winthrop		Tennessee	NC State	NC State	Tennessee
15	New Mexico St		Oklahoma	Temple	Colorado	NC State
15	North Dakota		Temple	Indiana St	Clemson	Temple
15	Jacksonville St		Colorado	Tennessee	Cincinnati	Colorado
15	Kent		Clemson	Clemson	Utah	Clemson

16	Troy		Arkansas	Oklahoma	Penn St	Texas
16	S Dakota St		Texas	Colorado	Arkansas	Arkansas
16	NC Central		Indiana St	Arkansas	UT Arlington	St Mary's CA
16	N Kentucky		UCF	Texas	Stanford	Indiana St
17	Mt St Mary's		Rhode Island	Rhode Island	Georgia	UCF
17	UC Davis		St Mary's CA	UCF	Boston College	Utah
17	TX Southern		California	VA Commonwealth	Texas	Rhode Island
18	New Orleans		Georgia	Dayton	Memphis	UT Arlington

Assumptions and Possible Errors

As with any manipulation and conversion of data, there are several assumptions and limitations to the findings presented in this write up. It should first be recognized that this approach deals strictly with the numerical side of a sport. It does not account for player injury, coaching changes, or unforeseen events that may occur in the world of athletics. On top of this, it is limited in its ability to generate a ranking that could be used to establish a predictable March Madness bracket. In the real-life tournament, teams receive spots based on winning their conference, the type of school they come from, and decisions from the conference committee. With this in mind, it is nearly impossible that our final ranking will align perfectly with the 68 teams chosen to compete in March Madness. The last assumption made during this experiment was that the four deviations to look at the data were the best possible approaches.

If this analysis was run again, there are several steps that can be taken to mitigate the possible errors within our findings. The first approach should be to look at other ways to weight the matrix in the PageRank algorithm. Differentiation like team popularity, the scoring history over the past 10 years, and funding received by the school may be underlying factors that indicate a higher likelihood of winning March Madness. We could also apply more weight to recent games to account for improvements over the course of a season. If we wished to focus solely on how the teams chosen for March Madness performed, we could do more research on how easy/difficult schedules during the tournament can factor into our algorithm. With more time, a cleaner, slightly better picture could be built of the college rankings.

Extensions

One aspect of March Madness that we do not account for in our algorithm that creates a discrepancy in our results versus the final bracket is that conference winners get an automatic spot in the bracket. Our algorithm simply ranks teams 1-68 regardless of conference. Although this causes us to select different teams, we propose that the teams we select represent the best 68 teams in the NCAA. Through the current selection process, there have been teams that are selected for the bracket who have won only about half of their games, while the teams selected for an at large bid may have lost only five of their 25 games in a season. Conference champions also may not be the best team in their conference. They only have to win the conference tournament, and their regular season record is not considered. Our algorithm could take into consideration the teams' regular season and postseason records to create a better ranking of all teams in Division I NCAA basketball. Despite the flaw that the true top 68 teams are not represented in the tournament, there may be merit to including these teams such as having a diverse range of teams compete and to give smaller teams the opportunity to play. If we wanted to include these teams in our bracket, we run our algorithm the same way, and we would simply select these teams first, and go down our ranking list and add the top 32 teams that were not already selected for the at-large bids.

Another way would be to take a more statistical approach. We could gather more data about the teams from past years to create a model of what a winning team looks like. We could look at variables that are currently in our model like number of wins and score difference and also add statistics about the players and teams like rebounds, field goals, free throw percentage, injuries, points per game, fouls per game, and difficulty of scheduled games. We could look at the number of top players leaving the team and the recruiting rank for the people coming onto the team. We could use these variables to create a power index where teams with a higher index have a higher ranking. The model would output the power index, and we would rank the teams according to their power index. We would use previous years results to test the model against official rankings to create an optimal power index rating.

Conclusion

College sports are one of the most lucrative enterprises in American industry. Raking in about a billion dollars every year, the NCAA provides entertainment to millions of viewers as well as a college education and athletic experience to over 450,000 athletes. It is with the large impact on society and the weight it carries in social culture that there exists a prominent need for NCAA to rank and represent teams fairly. With the PageRank algorithm, one can generate an unbiased list of college teams based on their scoring and past history. Although not a holistic view of the players, this concept provides a numerical based summary that is free of human error to a degree. By incorporating this approach into college sports across the nation the NCAA can provide an even playing field for all athletes.