

Ranking NCAA Men's Basketball Teams Using Weighted PageRank

Cuong Nguyen (University of Connecticut), Tiandong Wang (Texas A&M University),
Jun Yan (University of Connecticut), Panpan Zhang (University of Pennsylvania)

Abstract

Each year, the National Collegiate Athletic Association (NCAA) selects 64 college basketball teams for March Madness. We apply a recently developed algorithm, weighted PageRank (WPR), to rank the tournament joiners. We establish a network system for all the teams, and develop a practical mechanism quantifying the weights between all pairs of teams. Team records in previous years are used as prior information. To validate the results, we compare ours with NCAA selections, and adopt a LogLoss function to assess the performance of the rankings.

Related Research

The weighted scheme used in this research was inspired by the following papers:

Matthews, Nicole et. al. (2021). Application of PageRank algorithm to Division I NCAA men's basketball as bracket information and outcome predictive utility. *Journal of Sports Analytics* 7, 1–9.

Zhang, Panpan et. al (2021). PageRank centrality and algorithms for weighted, directed networks, *Physica A: Statistical Mechanics and its Applications*.

Problem

Challenge:

Every year, the NCAA committee selects the final 64 teams (plus 4 play-ins) out of over 350 teams in the regular season to compete in a final tournament called March Madness.

Such selections are sometimes subjective since they are done by a committee taking into account not only win/loss performance but also factors such as viewership, historical rivalries, popularity, etc.

Can we use another method to create more objective rankings? Weighted PageRank is an existing method with rooms for improvement.




Year	Number of Participants		Champion
	Regular Season	March Madness	
2017	351	64 (plus 4 play-ins)	 North Carolina
2018	351	64 (plus 4 play-ins)	 Villanova
2019	353	64 (plus 4 play-ins)	 Virginia

Table 1: An overview of March Madness selection process.

NCAA Method	Existing WPR Method	Proposed WPR Method
Use Win/Loss	Use <i>weighted</i> Win/Loss	Use <i>weighted</i> Win/Loss
Take into account subjective criteria	No prior	Use prior based on previous year's rankings
Finalized by a committee	Computerized rankings	Computerized rankings

Table 2: A quick comparison between NCAA selections and WPR methods

Method

Weighted Win Calculation

Weighted Win = Score Difference × Venue × Day

Score Difference: higher the larger the score difference

Venue: penalized for Home advantage, increased for Away

Day: higher the later the season

Prior Calculation

Prior = Average Rank from *previous year*

For example, in 2017, Villanova has rank 4 from ESPN Power Index, rank 8 from USA Today, and rank 2 from Power Rank system out of 351 teams in total. Then, in 2018, Villanova would have the prior of:

$\text{Prior}(\text{Villanova}) = 351 - (4 + 8 + 2) / 3 = 346.33$

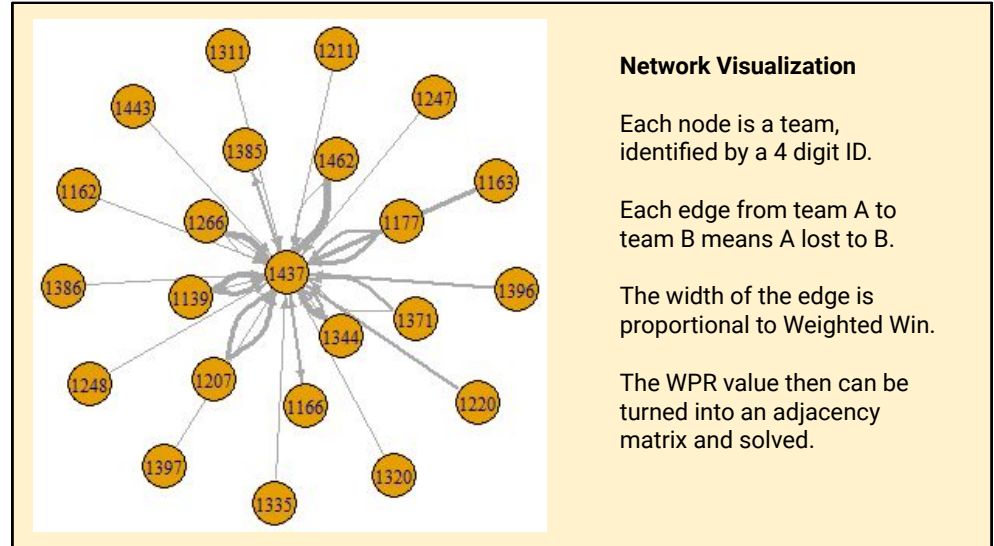
PageRank Calculation

$$\phi(i) = \gamma \sum_{j \in V} \left(\theta \frac{w_{ji}}{s_i^{(\text{out})}} + (1 - \theta) \frac{a_{ji}}{d_j^{(\text{out})}} \right) \phi(j) + \frac{(1 - \gamma) \beta_i}{\sum_{i \in V} \beta_i}$$

The implementation of WPR is available in R packet `wdnet` (gitlab.com/wdnetwork).

Winner	Loser	Score	Venue	Day (2018)	Weighted Win
UConn	Boston U	85 - 66	Home	20	3.75
UConn	Oregon	71 - 63	Neutral	24	4
Houston	UConn	81 - 71	Home	125	9.375

Table 3: Some examples of Weighted Win calculation.



Results

LogLoss Comparison (lower is better)

LogLoss is an objective function we want to minimize. More information about LogLoss can be found at kaggle.com/c/ncaam-march-mania-2021.

Generally, our method, which incorporates prior, outperforms the existing method, which does not.

Year	Existing WPR method LogLoss	Our WPR method LogLoss	Percentile compared to Kaggle Leaderboard
2016	0.62	0.58	60th
2017	0.59	0.56	30th
2018	0.62	0.60	71st
Average	0.61	0.58	

Table 5: LogLoss comparison between ours and existing method

NCAA Comparison

Applying our method to the 2018 tournament, we have found that 42 out of 64 teams selected by the NCAA for March Madness were in agreement. Villanova, the 2018 Champion, also had the highest WPR Score.

Team	WPR Rank	WPR Score	Actual Performance
Villanova	1	1.24	Champion
Michigan	2	1.03	Runner-up
W. Virginia	3	0.96	Top 16 (Eliminated by Villanova)
Virginia	4	0.95	Eliminated Round 1

Table 4: Overview of the top 4 teams ranked by our method

Conclusion

With high level of agreement with NCAA selections, as well as relatively good prediction power using LogLoss function, our method of using WPR with prior is a useful tool to select teams for March Madness.