Power-ful rankings of NCAA men's basketball teams

Hugo Hinckfuss, Josh Livingston, Magnus Miller (Section #002)

2024-04-29

The following paper discussed the derivation and development of multiple ranking schemes used to rank Men's NCAA College basketball teams in uneven paired competition (UEPC). This paper's motive to explore ranking schemes stems from the work of James P Keener who first used similar strategies in an application to rank college football teams. Through the direct use of the Perron-Frobenius theorem and numerical experiments to derive eigenvalues and eigenvectors, this paper developed a ranking scheme for uneven paired competition directly using the Perron-Frobenius theorem using a preference matrix comprised of wins and losses. This paper then expanded by considering other preference matrices to compare which include proportion of points scored and the square-root of points scored among others, eventually comparing them to the existing ranking methods such as the AP Coaches Poll. Results showed that limitations in the original UEPC system in terms of sparsity heavily hinder the ranking schemes that can be developed in this manner. Understanding this, however, by taking Keener's original work for ranking college football teams and applying the same methodology to college basketball, this paper has shown that with small alterations to the Perron-Frobenius theorem and through experimentation, similar ranking schemes can be developed for any other uneven paired competition based sport or event.

Attribution

Josh Livingston worked on the mathematical formulation used in the paper, developed and wrote the R code used to test the various ranking schemes, and handled the compilation of our work in Markdown. Hugo Hinckfuss wrote the introduction and worked on the mathematical formulation used throughout the paper, most notably in the development of the Direct Method ranking scheme. And Magnus Miller worked on the mathematical formulation, the abstract, and discussion as well as in editing the final paper.

Introduction

Every spring, the discourse surrounding collegiate basketball reaches its zenith as enthusiasts engage in discussions regarding the eventual supremacy of a particular team by the end of March. This discourse, however, is often clouded by subjective polls and rankings, notably those conducted by entities such as the Associated Press (AP), leading to confusion, disagreement, and frustration among fans who advocated for their respective teams. While certain publications propose alternative indices based on mathematical models, these methodologies often remain opaque to the general public due to the perceived complexity associated with mathematical concepts.

Our group's interest in the ranking of collegiate basketball teams peaked during this year's March Madness tournament, where several high-seeded teams succumbed to the unexpected triumphs of lower-seeded opponents. These upsets pushed our group to envision a future where a mathematically-driven ranking system could not only corroborate traditional voter-based polls but also enhance the precision of predicting teams' success in tournaments.

The difficulty in developing a ranking scheme in this type of system stems from the concept of uneven paired competition (UEPC), or comparison. In a UEPC based system, the outcomes of any paired comparison, or competition in the case of sports, is known but the full set of outcomes is incomplete. In the case of college basketball, this UEPC system has known outcomes (wins or losses) for a given game (comparison) throughout the season, but is incomplete as not every college team plays each other. In his paper, *The Perron-Frobenius Theorem and the Ranking of Football Teams,* James P Keener (Keener 1993) discovered that many ranking methodologies heavily leveraged the Perron-Frobenius theorem, which provided an elegant solution to the challenge of uneven paired competition.

This paper aimed to expand Keener's work to the world of NCAA Men's Basketball as it delved into various ranking methods, highlighting the utilization of the Perron-Frobenius theorem, eigenvalue computation techniques, fixed-point theorems, and probabilistic modeling. Each of the methods discussed in this paper formulate a ranking problem as a linear eigenvalue problem, directly applying the Perron-Frobenius theorem. Through various solutions to the entries of our matrix were we able to develop varying ranking schemes that prioritize different elements of college basketball to reveal the best.

This paper first develops an understanding of the Perron-Frobenius theorem through a pseudo-derivation of the theorem and expands on the mathematical formulation used throughout the development of the ranking schemes later introduced. The paper then builds on the Perron-Frobenius theorem to develop a direct ranking approach using wins and losses. Later the paper expands on this method deriving ranking priorities based on other factors of the game such as proportion of points scored and square root of points scored among others. Following the development of the ranking schemes, the paper analyzes the effectiveness of the proposed ranking schemes in comparison to the AP Top 25 poll and proposed hypotheses as to the successes and failures of the modeled schemes accuracy and validity.

Ultimately, we concluded by presenting the results derived from the application of these diverse ranking systems, aiming to showcase a collegiate basketball ranking that more accurately reflected the true capabilities of teams participating in March Madness. In the end, however, we discuss how the sparsity

of the data used in constructing our base matrices percolates throughout the formulation of our ranking schemes making it difficult to assess their viability. We propose, however, that due to the limitation of the original dataset, these ranking schemes may be better applied to NBA basketball where the UEPC is less uneven.

Mathematical Formulation

The Direct Method

This first method is regarded as the most direct method in ranking entities. In this method, we will simply assign a score to a team, depending on their result against an opposing team. The score a_{ij} given to the team will depend on the strength of the opponent, the margin of the scoring and the outcome of the game (win, loss or draw).

We will introduce the vector $\vec{r}$, which will represent the ranking values of teams, with positive component r_j representing the strength of the jth team (the opposition team). Then, the score for team i will be defined as

$$s_i = \frac{1}{n_i} \sum_{j=1}^{N} a_{ij} r_j$$

with a_{ij} being a nonnegative number depending on the aforementioned factors between teams i and j (Keener 1993). N will be the total number of teams in the system, which in our case will be larger than simply the teams that were allowed to play in March Madness, as we will require more data points to increase the rigidity of the standings. We will also have a focus on the PAC12 Conference, as it will be easier to see the effectiveness of this method relative to both their conference records, but also their AP Rankings and their ultimate performances in the March Madness tournament.

The matrix A with entries a_{ij} is referred to as a preference matrix. In our case, we will notate the entry of a_{ij} as 1 if team i is victorious, or 0 if the game is lost. There is also a column for draws, with an entry value of ½, but these will always be empty as there are no draws in collegiate basketball, yet this column will be valuable for use in other applications. The reason why the sum will be divided by n_{ij}, the number of matches between teams i and j, is to ensure teams cannot accumulate and artificially large score through playing more games than their competitors.

Next, the strength (which will be ordered to provide rank) of a participant will be a function of the team's score, given by

$$A\vec{r} = \lambda \vec{r}$$

where A is the matrix containing entries a_{ij}. Thus, the ranking vector $\vec{r}$ is a positive eigenvector from the positive matrix A (Keener 1993).

The Perron-Frobenius Theorem tells us that when the non-trivial matrix A has nonnegative entries, there exists an eigenvector $\vec{r}$ with nonnegative entries, which have a corresponding eigenvector λ. Furthermore, if A is irreducible, eigenvector $\vec{r}$ will have only unique and positive entries.

In order to calculate the eigenvector, we can take advantage of the power method - namely, the fact that the ranking vector $\vec{r}$ will correspond to the largest eigenvalue of A; i.e.:

$$\lim_{n \to \infty} \frac{A^n \vec{r}_0}{|A^n \vec{r}_0|} = \vec{r}$$

for an nonnegative vector $\vec{r}_0$.

Initially, with the system devised, all teams would be given an initial ranking of one. Then the ith component of the vector $A^2 \vec{r}_0$ would represent the average winning percentage of the teams defeated by team i (Wei 1952). Thus, this component of the vector would provide the reader with information regarding the strength of a teams schedule, and it is arguably a stronger metric than simply winning percentage. The predominant issue with this figure is that it places such a significant portion of the score on the strength of the schedule.

This system works really well for systems where teams play each other often during a season, and have different schedules i.e the NBA and MLB regular seasons (no team plays every team the same amount of times). In this case, a_{ij} will be the total victories of team i over j. Intuitively, with more and more games, this system will become a superior indicator of the relative strength of the two teams.

The issue arises in seasons of collegiate basketball and football, where teams will play each other very few times, often once. In this situation, an arbitrary score of 1,½ or 0 only provides information for the outcome. Thus, there will be lost information on how close the teams were in ability, as the same score is provided for a 120-60 win and an 120-1119 win. Furthermore, a team that is winless in a season will have rank zero and will therefore not contribute to their opponents score at all, which actually makes playing these teams harmful, or worse than not playing them at all, as it will only divide the numerator by a larger number.

In order to account for this, we can instead distribute a single point between the two teams in a continuous manner. This can be achieved through distributing the point according to the points scored by each team. Suppose team i scored s_{ij} points and team j scored s_{ji} points when they played each other, the input a_{ij} may be represented as

$$a_{ij} = \frac{s_{ij}}{s_{ij} + s_j}$$

for team i.

For application in other sports, particularly potentially low scoring sports such as football or soccer, the earlier problem arises if the losing team fails to score a single point/goal. In soccer, a score of 1-0 is

regarded as a tight result, yet the losing team would still be awarded 0. To account for this, we will assign

$$a_{ij} = \frac{s_{ij} + 1}{s_{ij} + s_{ji} + 2}$$

To further enforce the rigidity of this process, we also have to account for the fact that when a team is winning comfortably, they will often rotate weaker players or simply try less. In our system, the winning team showing 'mercy' simply weakens their potential score. To fix this problem, we also tried distributing the points in a non-linear manner. A function that we tried to use to more effectively rank teams was:

$$a_{ij} = h\left(\frac{s_{ij} + 1}{s_{ij} + s_{ji} + 2}\right)$$

$$h(x) = \frac{1}{2} + \frac{1}{2}s\left(x - \frac{1}{2}\right)\sqrt{|2x - 1|}$$

where $s(x)$ is the sign of x (Keener 1993).

Numerical Results

Data

We obtained team-game level data from sports-reference.com[1] for the 2023-24 NCAA men's regular season. After doing basic data preparation, we bundled the data in an R package in this report. The R package, including the data and code for the computation, figures, and tables, is available on Github[2].

A sample of the data is available below. In total we have a 10,654 x 7 dataset. There are two rows per game in, one row per team. And for each team-game thereare 7 columns: the team, date, home/away designation, opponent, result, team's score, and opponent's score.

Table 1: Sample of prepared NCAA men's basketball data

team	opp	result	team_score	opp_score
Texas State	Little Rock	L	66	71
Grand Canyon	Southeast Missouri State	W	88	67

---

[1]https://stathead.com/basketball/cbb/team-game-finder.cgi?request=1&comp_type=reg&game_status=1&order_by=date&match=team_game&year_max=2024&order_by_asc=1&timeframe=seasons&comp_id=NCAAM&year_min=2024

[2]https://github.com/joshwlivingston/appm3310.final

team	opp	result	team_score	opp_score
Mississippi Valley State	Louisiana State	L	60	106
Southern	TCU	L	75	108
Jackson State	Memphis	L	77	94
Pacific	Sam Houston State	L	57	64
Southern Utah	Cal State Bakersfield	L	72	73
Tarleton State	Virginia	L	50	80
Bethune-Cookman	Minnesota	L	60	80
UT Arlington	Oral Roberts	W	75	71

Eigenvector computation

Win-loss matrix

We will now explore in depth one method to compute the ranking eigenvector. In our case, we will notate the entry of a_{ij} as 1 if team i is victorious, or 0 if the game is lost. As there are no ties in college basketball, these two designations are sufficient for a win-loss matrix. A sample of this matrix is shown below. For simplicity, we'll let B_1 refer to the $4x4$ sample of A_1 with corresponding entries $b_{1ij} : i, j = 1, 2, 3, 4$.

Table 2: Sample of win-loss matrix

	Alabama	NC State	Purdue	Connecticut
NC State	0	0	0	0
Purdue	1	0	0	0
Alabama	0	0	0	0
Connecticut	0	0	0	0

In matrix notation, the matrix B is

$$B = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

The matrix entries $b_{1ij} \in B_1$ line up with team-game level results from the regular season.

Table 3: Regular season results for teams in the NCAA men's 2024 final four

team	date	home_away	opp	result	team_score	opp_score
Alabama	2023-12-09	NA	Purdue	L	86	92

A potential downside to this method, is that teams that have never matched have the same entry for b_{ij} as teams that have lost to another team. For both of this scenarios, $b_{ij} = 0$. You can see that $b_{33} = b_{32} = 0$, where $b_{33} = 0$ represents Alabama's loss to Purdue, and $b_{32} = 0$ represents Alabama and NC State never having played a game against each other in the regular season So, in this scenario, we lose information about losses. This can lead to teams with losses to have an inflated ranking.
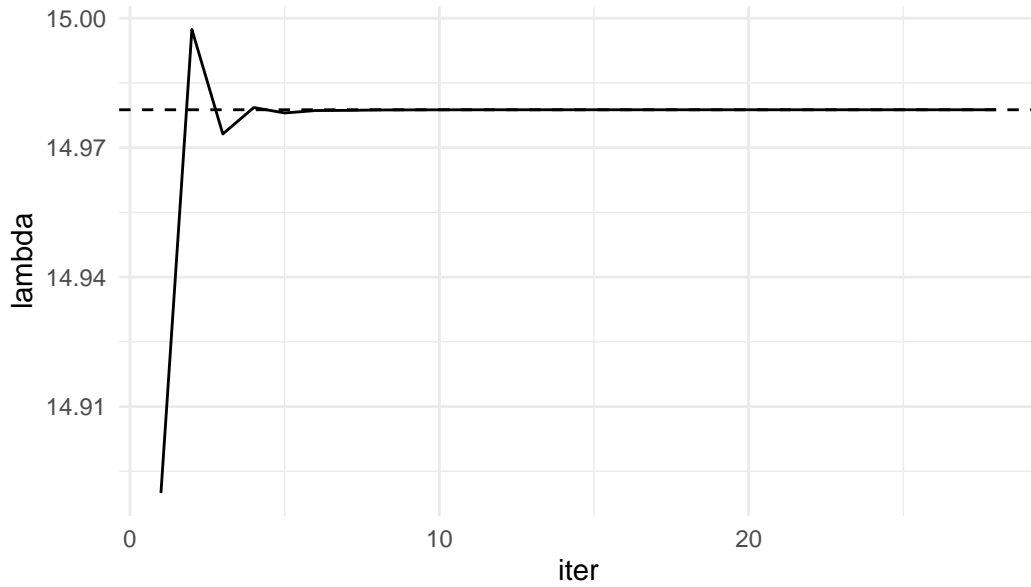
Eigenvalue approximation

We denote the eigenvalue λ for matrix A_1 as λ_1.

To approximate λ_1, we employ the power method. At each step of the power method, we compute the approximate eigenvalue λ_{1i} where i is the iteration step of the power method approximation, and

$$\lim_{i \to \infty} \lambda_{1i} = \lambda_1$$

After approximation, we observe λ_{1i} converging to λ_1, shown in the figure below.

Values of lambda converge to approximately 14.979 after 28 iterations.
Figure 1

We use the approximated eigenvalue to compute the corresponding approximate eigenvector. We can use this eigenvector to compute rankings for NCAA men's basketball teams following the 2023-24 season. The rankings are shown below for the AP top 25 teams, showing that the win-loss method is an ineffective ranking scheme for NCAA men's basketball.

Performance of ranking approximation

The win–loss ranking method is an ineffective ranking method.

Figure 2
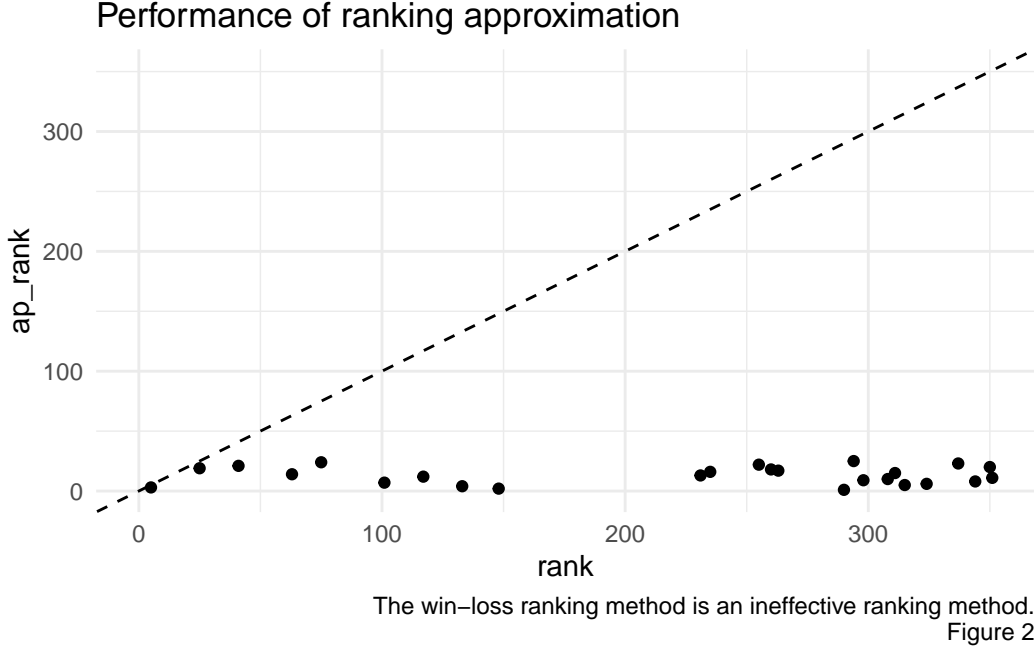
Additional Ranking Methods

Additional variations of a_{ij} did not produce better results. We used proportion of points, scored, various polynomial functions, distance between points scored, etc as values for a_{ij}.

To compare impact of values of a_{ij} on the resulting eigenvalues, we calculated the sum of squared difference between the eigenvector and the win-ratio. We denote this sum of squared difference for a matrix A_i $\vec{e}_i$ where i denotes the ranking function used to create the matrix A_i. Let $\vec{r}_i$ denote the ranking vector for A_i and $\vec{q}_i$ denote the win-ratio's for the teams in matrix A_i. That is,

$$\vec{q}_i = \frac{\vec{w}_i}{\vec{w}_i + \vec{l}_i}$$

where $\vec{w}_i$ represents the vector of teams' wins for a given ranking method in a given season and $\vec{l}_i$ represents the vector of teams' losses. That leads us to the following definition of $\vec{e}_i$:

$$\vec{e}_i = (\vec{r}_i - \vec{q}_i)^2$$

8

The best performing matrix is the A_i that minimizes $\vec{e}_i$. We denote this matrix A^*, with entries a_{ij}^*. The table below shows a summary of each A_i with its corresponding e_i.

Table 4: Table showing the top five ranking matrices with corresponding error metrics

ranking_table_name	e
points_proportion	84.121
points_polynomial5	84.151
points_sqrt_points	84.166
points_polynomial3	84.245
points_scored_mean	84.288

The best performing table is the table with entries a_{ij} constructed with the proportion of points scored in each game. That is,

$$a_{ij}^* = \frac{s_{ij}}{s_{ij} + s_{ji}}$$

The figure below shows the eigenvalues plotted against the win-ratio's for all 362 teams in the data using a_{ij}^* as the entries for A^*. We also included plotted the linear regression line between the two points, for reference. We can see that for the full league, the eigenvalues bear little relationship to the win-ratios, suggesting that none of the tested ranking methods are effective at ranking the entire NCAA.

Win rate vs approximated strength
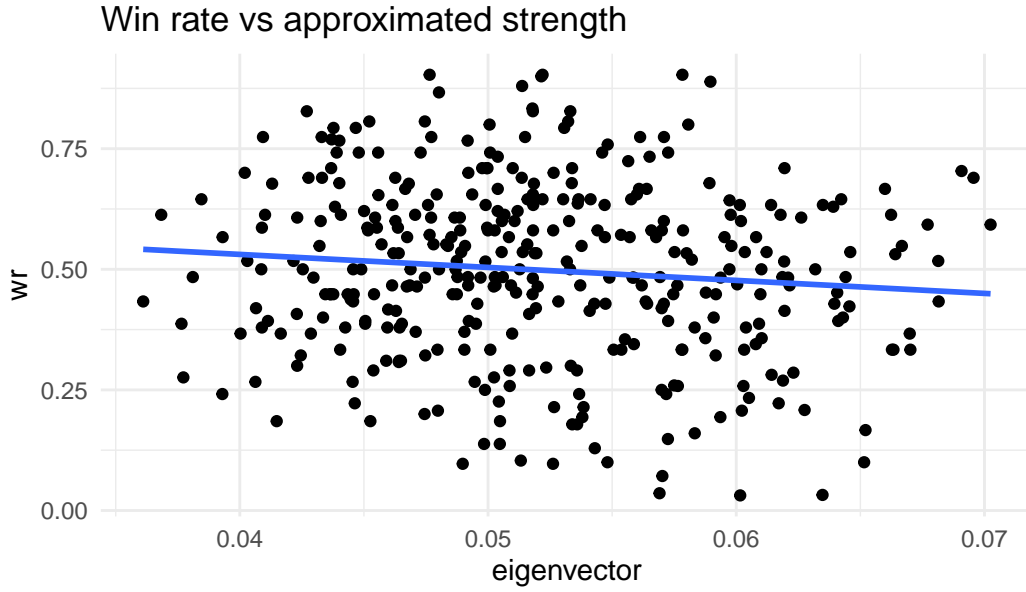
The top performing ranking scheme is ineffective at ranking the entire NCAA
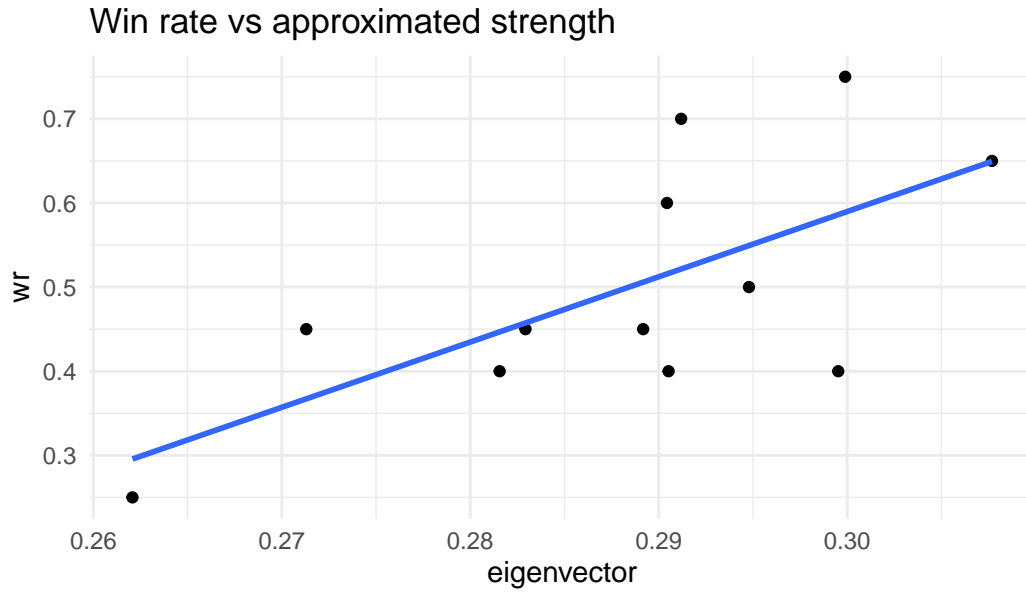Figure 3

Sparsity

One issue affecting results is the sparsity of the ranking matrices we are using. When analyzing the matrix comprised of 362 teams x 362 teams, most of whom do not face each other in the regular season, the matrix is going to be mostly sparse.

To explore the role of sparsity, we look at a second, filtered dataset, comprising of in-conference games between PAC-12 teams. That is, we only look at games in which two PAC-12 teams played each other. The figure below shows the same 5 ranking schemes are the most effective ranking schemes.

Table 5: Table showing the top five ranking matrices with corresponding error metrics

ranking_table_name	e
points_proportion	0.747
points_scored_mean	0.763
points_sqrt_points	0.764
points_polynomial5	0.776
points_polynomial3	0.782

Looking at the relationship between the eigenvector and the win-ratio, we see a much stronger relationship, suggesting sparsity limits the effectiveness of theme ranking schemes.

Win rate vs approximated strength

When limited to conference play, the ranking method via proportion is an effective ranking method.

Figure 4

The table below shows the ranking for PAC-12 using the proportion ranking matrix.

Table 6: Rankings for in-conference PAC-12 play for the NCAA men's 2023-24 basketball season as obtained via the proportion ranking matrix

team	eigenvector	rank	w	l	wr
Colorado	0.308	1	13	7	0.650
Arizona	0.300	2	15	5	0.750
Arizona State	0.300	3	8	12	0.400
UCLA	0.295	4	10	10	0.500
Washington State	0.291	5	14	6	0.700
Stanford	0.291	6	8	12	0.400
Oregon	0.290	7	12	8	0.600
California	0.289	8	9	11	0.450
Washington	0.283	9	9	11	0.450
Southern California	0.282	10	8	12	0.400
Utah	0.271	11	9	11	0.450
Oregon State	0.262	12	5	15	0.250

Discussion

The exploration of ranking methodologies for NCAA Men's Basketball has yielded insightful findings regarding the challenges and potential solutions in devising effective ranking systems moving forward. Our investigation was motivated by the prevalent subjective nature of traditional polls and rankings, which often lead to discrepancies and frustrations among fans and stakeholders especially during March Madness selection.

One of the central challenges addressed in this paper was the concept of uneven paired competition (UEPC), inherent in collegiate basketball due to a large number of participating teams and very short schedules (in comparison). This challenge faced by all ranking metrics, as identified by Keener in his work on college football teams, highlights the importance of leveraging mathematical tools like the Perron-Frobenius theorem to tackle the complexities of ranking systems in sports with incomplete data sets. Our exploration focused on tackling the UEPC issue in the world of NCAA Men's Basketball similar to how Keener formulated his ranking schemes using the Perron-Frobenius theorem. Keener's framework which uses the Perron-Frobenius theorem served as a great base to work from as it provides a way to directly calculate an ordered vector of teams from an original preference matrix which can be specified and defined based on elements of the system one is analyzing. Using basic linear algebra principles such as eigenvector and eigenvalue calculations, we developed a direct ranking approach. In the original attempt, the preference matrix was designed to contain win-loss information about head-to-head matchups which was later replaced by other metrics including proportion of points scored and square root of points scored, among others. These methodologies and ranking frameworks aimed to provide a more contextually based and nuanced formulation for ranking teams beyond the AP's method of considering wins and losses in conjunction with coaches' votes.

Our analysis revealed, however, the numerous limitations and considerations that need to be taken into account when developing a ranking scheme built on the Perron-Frobenius theorem, most notably being the sparsity of the original data. In the case of college basketball, there are simply too many teams that individually do not play enough games. This leads to gaps and missing information necessary for developing a ranking metric using the direct method derived in this paper. The sparsity of data, especially in non-conference matchups, posed challenges in accurately assessing team strengths across the entire NCAA landscape. This sparsity affected the accuracy and effectiveness of ranking schemes, as seen in Figure 2, which compares the Direct Method ranking and the AP ranking and Table 4 which outlines the error associated with the various Direct Method versions compared to the AP poll. In both cases, the data used stems from the analysis of the full NCAA men's basketball landscape. When these gaps are not considered, as in the case for PAC12 rankings, the sparsity issue disappears as each team plays the others at a more regular frequency as seen in Figure 4 and Table 6.

In conclusion, while our exploration into mathematical ranking methodologies based on the Perron-Frobenius theorem has provided valuable insights and alternative approaches to traditional polls, there are clear limitations to consider as the effectiveness of ranking schemes heavily depends on data availability and the nature of competition within the NCAA.

Moving forward, it is crucial for future researchers to address the issue of sparse data when it comes to

developing a direct method ranking scheme based on the Perron-Frobenius theorem as outlined in this paper. Investigations into the use of data augmentation techniques, other factors beyond records and points scored, dataset size, and models for nuanced elements of the game could be the key to developing a full and robust ranking scheme that competes with the AP Poll.

For practical application, our findings suggest that ranking methodologies, particularly those relying the Perron-Frobenius theorem, may find more success when applied to more structured and less sparse data sets such as. in-conference matchups within the NCAA basketball landscape or even in other leagues such as the NBA or MLB.

Ultimately, the pursuit of refining ranking systems in collegiate basketball requires a more multidisciplinary approach, combining mathematical rigor with domain-specific insights to address the nuances and challenges unique to the sport. In order to effectively derive a more fair and less opaque ranking scheme for men's collegiate basketball that could better suit the March Madness tournament selection format, continued research and refinement are necessary.

References

Keener, James P. (1993). "The Perron-Frobenius Theorem and the Ranking of Football Teams". In: *SIAM Review* 35, pp. 80–93. URL: http://www.jstor.org/stable/2132526 (visited on 04/29/2024).

Wei, Teh-Hsing (1952). "The algebraic foundations of ranking theory". In: University of Cambridge. URL: https://api.semanticscholar.org/CorpusID:58372584.