

Article

Applying PageRank to Team Ranking in Single-Elimination Tournaments: Evidence from Taiwan's High School Baseball

Yu-Chia Hsu *  and Wen-Jie Zhang

Department of Sport Information and Communication, National Taiwan University of Sport,
Taichung 404401, Taiwan

* Correspondence: ychsu@ntus.edu.tw

Abstract: This study examines the applicability of the well-established PageRank algorithm for ranking teams and predicting outcomes in incomplete, single-elimination high school baseball tournaments. Using match data from Taiwan's CTBC Black Panther Cup National High School Baseball Tournament spanning from 2013 to 2023, this research investigates whether PageRank can produce valid, stable, and predictive rankings under structural constraints and limited data environments. Three empirical evaluations were conducted. First, a comparative analysis between PageRank rankings and official results demonstrated high ordinal consistency, with Kendall's tau values exceeding 0.70 in most seasons. Second, PageRank rankings were assessed for temporal robustness, demonstrating stable performance across seasons and under varying data inputs. Third, a series of n-step-ahead simulations were implemented to test the predictive validity of PageRank. The results indicate that incorporating historical data substantially improves forecasting accuracy, achieving up to 92.9% when data from up to four previous seasons are included. Overall, the findings support PageRank as a consistent and interpretable ranking method that is well-suited for grassroots sports. Its ability to incorporate indirect competition and opponent strength makes it effective in settings with sparse or unbalanced schedules. This study provides methodological insights and practical implications for ranking and evaluation in school-level sports.



Academic Editors: Vasileios Armatas
and Sotirios Drikos

Received: 19 May 2025

Revised: 10 June 2025

Accepted: 12 June 2025

Published: 18 June 2025

Citation: Hsu, Y.-C.; Zhang, W.-J. Applying PageRank to Team Ranking in Single-Elimination Tournaments: Evidence from Taiwan's High School Baseball. *Appl. Sci.* **2025**, *15*, 6882. <https://doi.org/10.3390/app15126882>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: sports ranking; match outcome prediction; sports analytics; incomplete competitions; PageRank algorithm

1. Introduction

In competitive sports, ranking systems and tournament structures play a critical role in ensuring the procedural legitimacy of outcomes. Whether used for playoff qualification, seed allocation, or athlete selection, rankings are expected to reflect actual team or player performance while minimizing the influence of random variation and structural imbalance. When a ranking system fails to align with observed performance, it can undermine the transparency, credibility, and predictability of the competition framework. The inaccurate evaluation of team strength distorts the interpretation of comparative performance and limits the utility of rankings as a basis for competition outcomes [1]. Particularly in team sports at the school level, the robustness and consistency of ranking models carry broader implications, influencing regional athletic development, resource allocation, and the design of sports governance policies [2,3].

In most contemporary baseball tournaments, team rankings are typically determined by statistical indicators, such as the number of wins, win percentage (win rate), or run

differential. These methods are widely adopted across various levels of competition due to their computational simplicity, interpretability, and institutional transparency [4]. Among them, the most commonly used approach is win percentage, which calculates a team's success rate based on the proportion of games won. This method is intuitive and operationally straightforward, offering a basic performance metric. However, it critically depends on a complete and balanced schedule. When the competition structure is highly incomplete or asymmetric, the win percentage fails to capture qualitative differences in the opponents faced and outcomes achieved [5]. From a representational standpoint, teams with stronger opponents may be penalized in the rankings, while those with weaker schedules may receive disproportionate advantages.

Another traditional method is run differential, which quantifies overall team performance by subtracting runs allowed from runs scored across games [6]. Although this approach provides a finer-grained metric of offensive and defensive capability, it is highly sensitive to outliers, such as one-sided games or unpredictable results. In short seasons or elimination-based formats, these fluctuations may distort rankings and place resource-constrained teams, which often lack the opportunity to accumulate consistent long-term results, at a disadvantage.

To address the limitations of purely statistical ranking systems, mathematical models such as the Colley and Massey methods have been proposed. These models incorporate opponent strength to mitigate schedule imbalance. The Colley matrix uses systems of linear equations based on pairwise outcomes to generate strength estimates and is noted for its transparency and objectivity [7]. The Massey method minimizes the squared error between observed and predicted score margins, aiming to produce rankings that reflect relative performance more equitably [8]. While both methods have demonstrated efficacy in professional or collegiate baseball leagues with long and relatively complete seasons, their application in high school tournaments or single-elimination settings remains limited due to sparse data and potential ranking distortions [9–11].

Although these traditional methods remain widely accepted in practice for their simplicity and intuitive logic, they often struggle to account for latent performance differentials and structural biases under incomplete schedules or highly uneven matchups. Thus, there is a growing need to develop or adapt ranking models that can generate more representative and robust outcomes under real-world tournament constraints. The empirical evaluation of such models in authentic competitive environments remains an important direction for further research.

This challenge is especially acute in school-level competitions, where structural limitations are more severe. For example, Taiwan's largest high school baseball tournament—the CTBC Black Panther Cup National High School Baseball Tournament (hereafter referred to as the Black Panther Cup)—draws nearly 200 teams annually. Due to its scale and logistical constraints, the tournament adopts a single-elimination format. While efficient, this structure introduces substantial difficulties in evaluating team performance and constructing representative rankings. Matchups are highly unbalanced, and strong teams may be prematurely eliminated due to unfavorable draws. Traditional ranking criteria based on win totals or win percentage fail to account for the relative difficulty of opponents, treating all victories as equivalent regardless of competitive context.

This study evaluates the applicability of the PageRank algorithm in modeling team strength within such constrained environments. While PageRank has been applied in various sports contexts [12–14], prior research has primarily focused on predictive performance in professional settings, with limited attention to its capacity to generate representative rankings under structurally imbalanced school-level competitions.

To address this gap, we analyze data from Taiwan's Black Panther Cup (2013–2023), applying the PageRank algorithm to model team strength under incomplete, single-elimination schedules. This study aims to evaluate the following: (1) the validity of PageRank in reflecting relative performance; (2) its consistency with official rankings and predictive stability over time; (3) its feasibility as a low-data, system-level tool for grassroots sports evaluation. The findings aim to inform tournament design and policy development through a performance evaluation framework that does not depend on individual statistics.

To address this gap, this study analyzes data from Taiwan's Black Panther Cup (2013–2023), applying the PageRank algorithm to estimate team strength under incomplete, single-elimination schedules. The objectives of this research are threefold: (1) to assess the validity of PageRank in capturing relative team performance; (2) to examine its consistency with official rankings across seasons; and (3) to evaluate its feasibility as a low-data tool for predicting match outcomes by leveraging historical data in grassroots sports contexts. The findings aim to inform tournament design and policy development by providing an alternative evaluation framework that does not rely on individual-level statistics.

2. Literature Review

2.1. Sports Ranking Algorithms

Research on sports ranking systems has long focused on producing consistent and adaptive assessments of team strength. Elo's rating system [15], initially designed for chess, introduced a dynamic framework that updates scores based on match outcomes to reflect relative skill. Glickman extended this by introducing the Glicko system [16], incorporating rating reliability to better capture short-term performance fluctuations.

Another major line of development involves probabilistic models that estimate win probabilities based on latent team strength. The Bradley–Terry model [17] laid the probabilistic foundation for pairwise comparisons, while the Keener ranking [18] applied steady-state distributions from Markov chains to model relative team strength. This approach has influenced many modern network-based ranking methods.

Several models have also incorporated score differentials to capture performance more precisely. Massey's least-squares method [8], for example, minimizes the difference between predicted and observed score margins, offering a more fine-grained representation of team ability.

With the rise of network science, researchers began treating competition results as directed graphs, where teams are nodes and game outcomes define edges. Govan et al. [12] proposed a PageRank-based model that uses this structure to propagate reputational value through indirect wins, demonstrating greater stability and representativeness in incomplete competition networks.

2.2. Incomplete Competitions and Ranking Issues

In practice, the robustness of ranking systems is challenged by the structural features of sports competitions. One issue is cyclic dominance, where non-transitive outcomes (e.g., A beats B, B beats C, but C beats A) are common. Park and Newman [19] found that these cycles accounted for 25–30% of outcomes in American college football, complicating linear ranking assumptions.

Another issue arises from the time dynamics of team performance. Stefani [20] noted that most international sports rating systems fail to properly address issues of performance decay or aging, leading to historical results having an outsized influence on current rankings. To address this, Barrow et al. [21] conducted an empirical comparison of eight commonly used sports ranking methods and recommended incorporating a time-weighted

win percentage to better capture recent performance and improve the temporal relevance of ranking outcomes.

While many academic ranking models have shown success in professional and international settings with near-complete schedules, their assumptions are difficult to satisfy in school-level tournaments constrained by limited resources and large team counts. This discrepancy is particularly evident in the case of Taiwan's Black Panther Cup, where the single-elimination format determines the first- and second-place teams solely based on the final match result. For teams eliminated in the same knockout round, specifically those ranked third through eighth, final standings are further separated using secondary performance indicators. These include the scoring rate, calculated as the total runs scored per inning or game, and the runs-allowed rate, which serves as a tiebreaker. Such efficiency-based metrics provide a pragmatic yet limited framework for distinguishing between teams with identical elimination outcomes, underscoring the practical challenges of producing meaningful rankings under structurally incomplete formats.

2.3. PageRank in Sports

The PageRank algorithm, developed by Brin and Page for web search applications [22], is grounded in Markov chain theory and random walk models. It calculates the importance of a node based on both the quantity and quality of incoming links. Beyond web search, PageRank has been widely used in fields such as citation analysis, social network analysis, and recommender systems.

The algorithm's logic transfers naturally to sports settings: Teams are modeled as nodes, and match outcomes are represented as directed edges pointing from the loser to the winner. A victory not only confers a direct score but also inherits part of the defeated team's strength, thereby reflecting both direct and indirect competitive performance. Empirical applications of PageRank have demonstrated its promise across a range of sports.

In professional sports settings where comprehensive match data are available, PageRank has been widely employed in research on predicting game outcomes. Regarding basketball, Shi and Tian [23] applied a generalized PageRank model to National Basketball Association data to enhance predictive accuracy in sports rankings. Their method constructs a game graph from a Bayesian correction perspective, representing teams as nodes and assigning edge weights via a link function based on match outcomes. Parameters are optimized by minimizing the loss between predicted and actual rankings, with results showing improved performance over conventional ranking models. In baseball, Yun et al. [24] analyzed 4861 Major League Baseball games from the 2017 and 2018 seasons. Their results demonstrated that PageRank provided stronger predictive validity than traditional win-loss-based ranking systems.

When individual-level statistics are unavailable, network-based ranking approaches that rely on aggregated competition data, such as win-loss records or cumulative points, provide a practical alternative. PageRank has demonstrated reliable performance under such low-data conditions. As reported in [25], the PageRank rankings for male 100 m sprinters in the 2016 IAAF Diamond League exactly matched the top five positions in the official points system and outperformed other ranking models, including Keener, Colley, and Massey.

Compared to methods based on win percentage or score aggregation, PageRank is better suited for handling unbalanced schedules, indirect match relationships, and non-round-robin formats such as single elimination. In the context of Olympic Taekwondo, Kim and Jeon [26] and Oh et al. [27] demonstrated the applicability of the PageRank algorithm for athlete ranking. Kim and Jeon [26] analyzed 144 matches in the Men's -58 kg category and found that PageRank produced rankings that were more valid than those from the

official Accumulative Point Index System (APIS). Similarly, Oh et al. [27], using 624 matches from the 2015 to 2018 World Grand Prix series, showed that PageRank outperformed the APIS by better accounting for opponent strength and enhancing ranking interpretability.

Continued research efforts have sought to improve the application of PageRank in sports ranking contexts. A recent refinement of PageRank, termed bi-directional PageRank, was proposed and tested on synthetic datasets calibrated from 18 empirical sports leagues such as soccer, basketball, baseball, and ice hockey [28]. These datasets were generated using a two-parameter variant of the Bradley–Terry model that accounts for home advantage and sport-specific randomness. The study found that standard PageRank outperforms the win ratio only in the early stages of a season and primarily in sports with low randomness. In contrast, bi-directional PageRank consistently demonstrated higher predictive accuracy by incorporating the quality of both victories and defeats.

To further enhance ranking precision, Lee et al. [29] proposed an artificial-intelligence-enhanced version of the PageRank algorithm. This model was applied to 1466 NCAA Division I football games, with the damping factor adaptively tuned to reflect the specific characteristics of each dataset. This adaptive model achieved higher predictive accuracy (71.6%) compared to traditional PageRank (65.7%) and win-percentage methods (64.2%), while also demonstrating improved responsiveness to performance fluctuations and intra-conference dynamics.

Overall, PageRank provides a low-data, match-outcome-based approach that is particularly suitable for interscholastic competitions characterized by limited information and unbalanced schedules. Its reliability, consistency, and predictive capabilities have gained increasing recognition in the field of sports analytics.

3. Methodology

3.1. Research Design

This study employs an empirical framework structured in three stages to evaluate the applicability and stability of the PageRank algorithm in ranking teams under incomplete, single-elimination tournament structures. First, a competition network is constructed from historical match outcomes, and PageRank scores are computed to estimate relative team strength. Second, the consistency and temporal stability of the resulting rankings are assessed using statistical correlation metrics across multiple seasons. Third, the predictive validity of the rankings is examined through match-level simulations, where outcomes are forecasted based on pre-match PageRank scores and compared against actual results. The analysis further explores the effect of varying historical data horizons to compute PageRank relative to prediction accuracy.

3.2. Data Collection and Processing

This study utilizes data from the CTBC Black Panther Cup National High School Baseball Tournament in Taiwan, covering the period from its inaugural season in 2013 to 2023. The dataset comprises comprehensive match records from the preliminary rounds to the championship finals, totaling 1914 games. Data were obtained primarily from the official website of the Black Panther Cup (<https://blackpanthercup.tw/>, accessed on 4 May 2025) and supplemented by earlier historical records provided by the Chinese Taipei Baseball Association (<http://www.ctba.org.tw/>, accessed on 4 May 2025). These sources do not indicate explicit license restrictions for academic use, and all data were utilized exclusively for non-commercial academic research purposes in accordance with fair use principles.

The collected records were processed using standard data-cleaning procedures, including the removal of duplicate entries, imputation of missing values, and correction of inconsistencies to ensure overall data integrity. The cleaned data were then organized into a

structured database format to facilitate subsequent analysis. The finalized dataset includes key variables, such as tournament-stage classifications, match schedules, team names for each matchup, game scores, and match outcomes.

3.3. PageRank Modeling

This study employs the PageRank algorithm [22] to model team strength. As detailed in Section 2.3, its network-based approach is well-suited for sports. For this research on single-elimination high school tournaments, PageRank was chosen for its ability to handle sparse data and unbalanced schedules and qualitatively assess victories by considering opponent strength. Team rankings are calculated recursively from match outcomes in a directed network.

The PageRank score $r(i)$ for team i is computed as follows:

$$r(i) = \alpha \sum_{j \in B_i} \frac{r(j)}{N_j} + \frac{1 - \alpha}{n} \quad (1)$$

where B_i is the set of teams defeated by team i , N_j is the number of teams defeated by team j , n represents the total number of teams, and α denotes the damping factor, conventionally set at 0.85. This formulation ensures that victories against stronger opponents contribute more significantly to a team's final ranking, capturing both direct and indirect relationships within the tournament structure.

For implementation in this study, match outcomes from the dataset were first structured into a directed competition network, with teams as nodes and match results represented as edges directed from the losing team to the winning team. Using this network, a transition probability matrix based on a Markov chain model was constructed to represent team interactions. The SciPy library (v1.13.1) under Python (v3.12.3) was subsequently employed to perform eigenvector computations via the power iteration method, thereby deriving the steady-state distribution of the transition matrix and calculating PageRank scores for each team. These scores served as the foundation for both team ranking and match outcome prediction in the subsequent stages of analysis.

3.4. Evaluation Procedures

To evaluate the performance of the PageRank rankings, three empirical tests were conducted:

- Ranking Consistency Analysis

To quantify the ordinal agreement between the PageRank output and the official tournament standings, Kendall's tau rank correlation coefficient (hereafter, Kendall's tau) was employed. As an established and widely adopted non-parametric metric, Kendall's tau remains highly suitable for this analysis because it effectively captures pairwise ranking reversals and provides a robust indicator of structural similarity between rankings. Higher Kendall's tau values indicate stronger consistency, with values approaching 1 reflecting near-perfect concordance. The coefficient is calculated as follows:

$$\tau = \frac{C - D}{\frac{1}{2}n(n - 1)} \quad (2)$$

where C and D are the numbers of concordant and discordant pairs among n teams. Statistical significance is assessed via a two-tailed p -value, with $p < 0.05$ indicating a meaningful correlation.

- Simulation-Based Prediction

Forecast accuracy is tested by simulating each season's matches. For every game, the team with the higher pre-match PageRank score is predicted to win. Forecast accuracy is then computed as follows:

$$\text{Accuracy} = \frac{\text{Number of correctly predicted matches}}{\text{Total number of matches}} \quad (3)$$

This metric is calculated on a per-season basis to assess the model's out-of-sample predictive reliability.

To evaluate the predictive utility of PageRank in estimating team strength, a series of empirical tests was conducted using historical match data. An n -step-ahead prediction framework was implemented, in which PageRank scores were calculated prior to each game based on match results from the current season and up to five preceding seasons.

Each tournament season follows a single-elimination format comprising four stages: preliminary (P), quarterfinals (Q), semifinals (S), and final (F). After the preliminary round, eight teams advance to four quarterfinal matchups, followed by two semifinal games (three in 2013 and 2014 due to format variations) and culminating in the championship final. To illustrate the n -step-ahead prediction framework, a one-year historical window was employed to forecast the 2015 quarterfinal outcomes. PageRank scores were computed using all matches from the 2014 season (including the preliminary, quarterfinal, semifinal, and final stages), along with the 2015 preliminary results. Similarly, to predict the semifinals, both the complete 2014 season data and the 2015 preliminary and quarterfinal matches were used to estimate PageRank scores prior to those games.

- **Temporal Robustness Assessment**

To evaluate temporal robustness, two complementary analyses were performed using data from the 2013 to 2023 seasons. First, Kendall's tau was calculated between PageRank rankings and official standings in consecutive years to assess the year-to-year consistency of estimated team strengths. Second, predictive accuracy was evaluated across different historical data horizons, as defined in the simulation framework. These assessments demonstrate the stability and reliability of the PageRank rankings across tournament seasons.

4. Empirical Results and Discussion

4.1. Overview of PageRank Rankings

Based on data from the 2013 to 2023 tournaments, average PageRank scores were computed for all teams across seasons. Figure 1 presents the top 10 teams by their average PageRank, along with their corresponding win percentages (win rate), which are defined as wins divided by the total games played.

Teams such as Pingzhen High School (HS), Kupao Home Economics and Commerce HS, and Kaoyuan Vocational HS consistently ranked among the top three across both metrics. Their PageRank values more clearly distinguish them from other teams, reflecting sustained competitive dominance. Notably, Pingzhen HS leads with a long-term average PageRank of 0.609 and secured the most Black Panther Cup titles over the study period.

Several teams exhibit notable discrepancies between their PageRank and win-rate rankings. For instance, Taoyuan Agricultural & Industrial HS ranks fourth in PageRank but fifth by win rate, suggesting stronger opponents defeated and a higher performance value that is not reflected in the raw win percentage. In contrast, Yingge Vocational HS ranks higher by win rate but lower in PageRank, indicating that their success may have come against weaker teams.

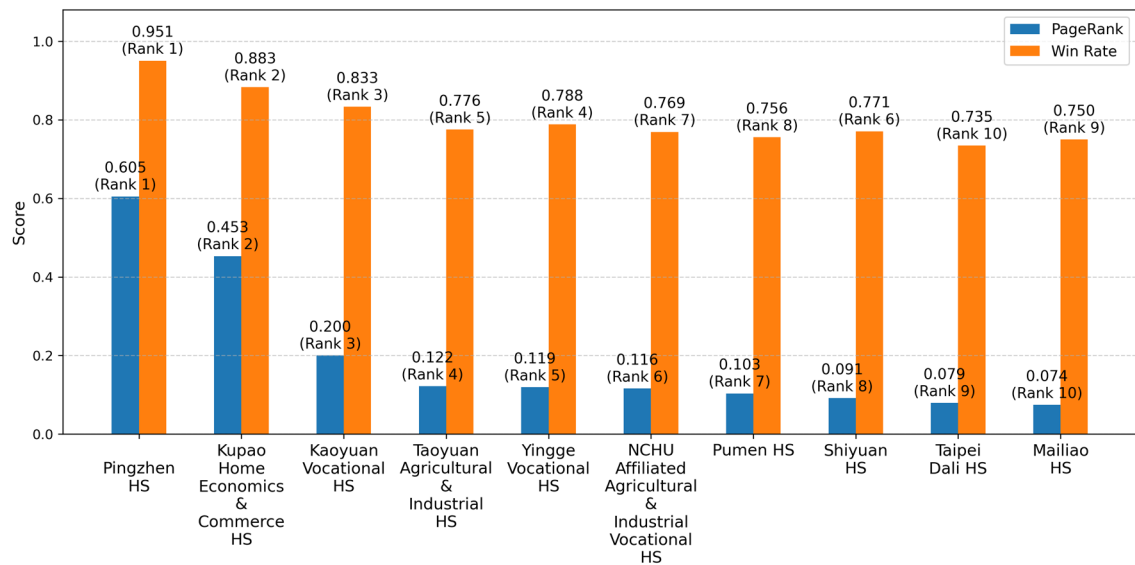


Figure 1. Comparison of the top 10 teams by average PageRank score and win rate (2013–2023).

Similarly, Shiyuan HS ranks 6th in PageRank but only 10th in win rate, likely due to impactful wins over high-strength opponents despite a modest overall record. Conversely, Mailiao HS ranks 6th in win rate but 10th in PageRank, reflecting inflated results from favorable matchups. These cases demonstrate PageRank’s capacity to capture performance quality beyond win–loss totals, offering a more nuanced measure of team strength in structurally imbalanced tournaments.

To further analyze longitudinal performance patterns, all 17 teams that reached the quarterfinals at least once during the 2013–2023 seasons were identified, and their annual PageRank scores were visualized in a heatmap (see Figure 2).

The heatmap illustrates a clear contrast between teams exhibiting sustained excellence and those demonstrating episodic success. Teams such as Pingzhen HS, Kupao Home Economics & Commerce HS, and Kaoyuan Vocational HS consistently display high PageRank values across multiple seasons, underscoring their long-term dominance and structural centrality within the tournament network. These stable patterns reflect not only frequent advancement into later rounds but also victories over strong opponents, which PageRank effectively captures through its recursive weighting mechanism.

Conversely, teams such as Pumen HS, Yingge Vocational HS, and NCHU Affiliated Agricultural & Industrial Vocational HS show marked spikes in PageRank scores during isolated years, suggesting that transient competitive surges are likely tied to cohort effects, favorable draws, or exceptional individual players. This episodic behavior contrasts with their generally lower presence in other seasons, indicating weaker overall program stability.

Furthermore, the matrix suggests that only about 3 to 4 teams consistently attain high PageRank scores across multiple years, while most other teams remain low-ranked. This concentration indicates a relatively uneven competitive landscape, with dominance clustered among a few programs. Nevertheless, the occasional emergence of 6 to 8 different teams reaching high scores in select years also reflects a degree of mobility and turnover, characteristic of youth-level tournaments where talent development and cohort variability play significant roles.

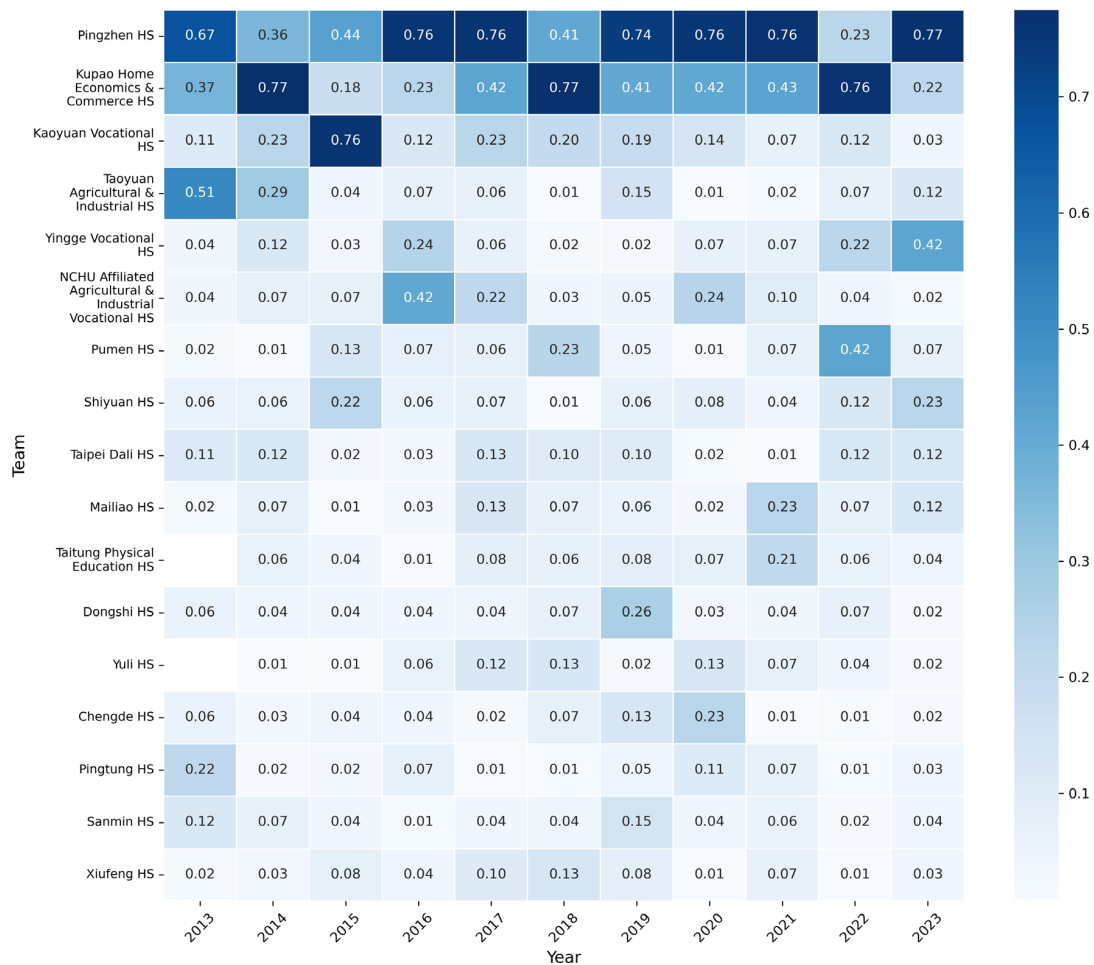


Figure 2. PageRank heatmap by team and year.

4.2. Ranking Consistency Analysis

Kendall’s tau rank correlation coefficient was employed to evaluate the relationship between PageRank rankings and the official final standings. Using the top eight teams recognized by the tournament’s official awards each year as the reference sample, Kendall’s tau was computed for each season from 2013 to 2023, as shown in Table 1. In 10 out of the 11 seasons, Kendall’s tau values exceeded 0.70, and in most cases, the corresponding *p*-values were below 0.05, indicating statistically significant correlations. Notably, the values for 2014 and 2018 reached 0.929 and 0.905, respectively, suggesting near-perfect agreement between PageRank rankings and official results. The lowest correlation occurred in 2019, with a Kendall’s tau of 0.643.

Table 1. Kendall’s Tau correlation and significance (2013–2023).

Year	Kendall’s τ	<i>p</i> -Value
2013	0.786 **	0.0055
2014	0.929 ***	<0.001
2015	0.733	0.0556
2016	0.810 *	0.0107
2017	0.786 **	0.0055
2018	0.905 **	0.0028
2019	0.643 *	0.0312
2020	0.786 **	0.0055
2021	0.714 *	0.0141
2022	0.714 *	0.0141
2023	0.810 *	0.0107

Note: * *p* < 0.05; ** *p* < 0.01; *** *p* < 0.001.

Overall, the consistently high Kendall's tau values demonstrate the reliability and validity of PageRank as an objective ranking method. In particular, the method proves to be useful in single-elimination formats where official rankings are not assigned beyond the top eight, offering a robust and interpretable estimate of relative team strength in the lower ranks as well.

4.3. Match Outcome Prediction

The predictive performance of PageRank rankings was evaluated using six temporal horizons, ranging from the current season alone to the inclusion of up to five preceding seasons. Table 2 summarizes the results. When relying solely on current-season data, prediction accuracy was low (41.8%). In contrast, incorporating one to five years of prior results significantly improved performance, with accuracy ranging from 76.2% to 83.7%.

Table 2. Prediction accuracy by historical data horizons and tournament stages.

Horizon (Years)	Stage	Total Games	Correct Predictions	Accuracy
0 (current season)	All (Q, S, F)	79	33	0.418
	Quarterfinals	44	21	0.477
	Semifinals	24	10	0.417
	Finals	11	2	0.182
1	All (Q, S, F)	71	55	0.775
	Quarterfinals	40	34	0.85
	Semifinals	21	14	0.667
	Finals	10	7	0.7
2	All (Q, S, F)	63	50	0.794
	Quarterfinals	36	31	0.861
	Semifinals	18	13	0.722
	Finals	9	6	0.667
3	All (Q, S, F)	56	44	0.786
	Quarterfinals	32	25	0.781
	Semifinals	16	13	0.812
	Finals	8	6	0.75
4	All (Q, S, F)	49	41	0.837
	Quarterfinals	28	22	0.786
	Semifinals	14	13	0.929
	Finals	7	6	0.857
5	All (Q, S, F)	42	32	0.762
	Quarterfinals	24	17	0.708
	Semifinals	12	10	0.833
	Finals	6	5	0.833

Figure 3 illustrates how accuracy varied by tournament stage. For quarterfinal predictions, using data from the current season and two previous years yielded the best results. Semifinal predictions required a broader historical window of three to five years to achieve optimal accuracy. Moreover, we observed that in high-stakes rounds such as the semifinals and final, predictive accuracy consistently increased with longer historical input.

The decline in accuracy when using 5-year historical data (76.2%) compared to 4-year data (83.7%) reflects several external influences that affect the tournament's competitive dynamics over time. These include tournament format changes (such as the additional semifinal match in 2013–2014 seasons and the elimination of seeded bracket protections in 2020), the temporal decay of information due to natural program evolution and player turnover cycles typical in high school sports, and COVID-19 impacts on the 2020–2021

tournaments, where key players were unable to participate due to infection or quarantine protocols. Conversely, the exceptionally high Kendall coefficients in 2014 (0.929) and 2018 (0.905) suggest periods of greater competitive stability, where PageRank rankings aligned particularly well with tournament outcomes, possibly indicating more predictable competitive hierarchies during these seasons.

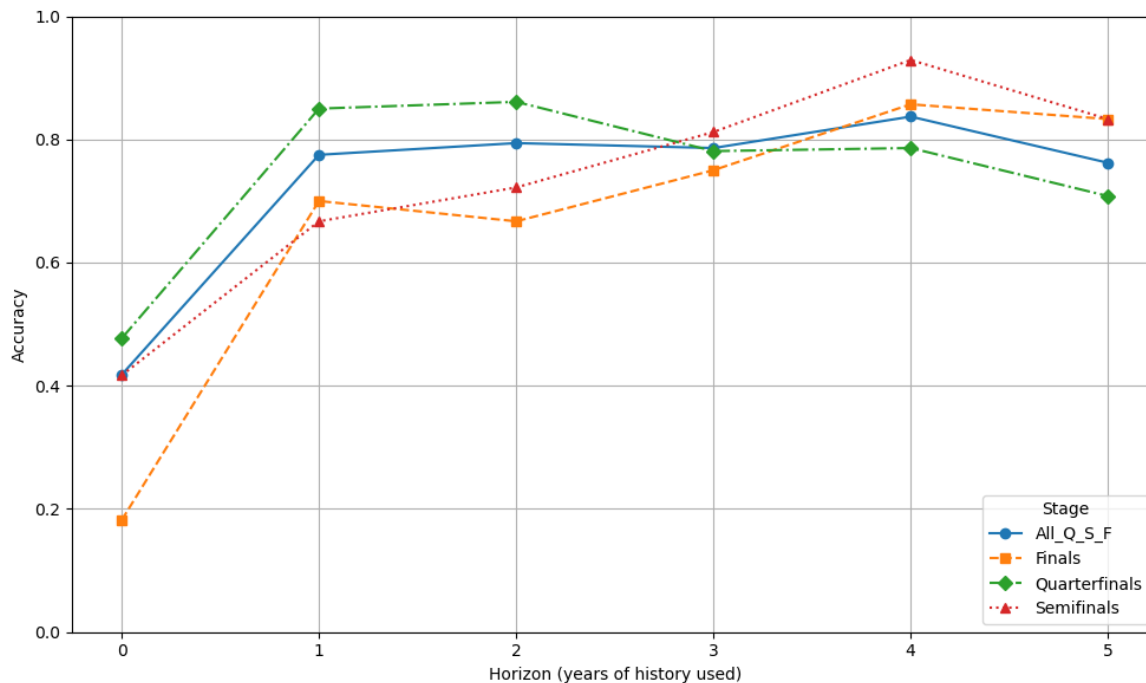


Figure 3. Prediction accuracy by horizon and stage.

This analysis suggests that 3–4 years represents the optimal balance between historical depth and temporal relevance for PageRank-based predictions in this context. The deterioration in predictive performance beyond this window indicates that older data may introduce noise rather than signals, particularly when structural changes to the tournament format or external disruptions alter the underlying competitive dynamics. In the later knockout stages, where competition is more selective and match outcomes are harder to infer, multi-year performance data provides more reliable signals of underlying team strength. The inclusion of longitudinal data mitigates the sparsity inherent in single-elimination structures and enhances the model’s ability to distinguish relative competitiveness. This underscores the importance of historical depth when applying network-based ranking methods to incomplete or asymmetric competitions.

5. Conclusions and Suggestions

5.1. Summary of Findings

This study examined the utility of the PageRank algorithm as a ranking and predictive tool in high school baseball tournaments characterized by incomplete and single-elimination formats. Drawing on a decade of empirical data from Taiwan’s Black Panther Cup (2013–2023), this study assessed whether PageRank can validly estimate team strength, align with official results, and offer forecasting capabilities using minimal input data.

The findings show that PageRank successfully identified top-performing teams and provided rankings highly consistent with official outcomes, as evidenced by strong Kendall’s tau correlations in 10 out of 11 seasons. Furthermore, the model yielded credible match predictions, with accuracy reaching as high as 92.9% when historical data from preceding seasons were incorporated.

While these results are derived from a single tournament context, they provide valuable insights into PageRank's behavior under structurally constrained conditions typical of school-level competitions. The case-based approach allows for the detailed examination of algorithmic performance within a specific competitive framework, offering methodological contributions that may inform applications in similar educational or resource-limited settings.

Overall, the findings support PageRank as a consistent and data-efficient approach for ranking and forecasting in incomplete or asymmetric competitions. Its low data requirements and structural flexibility make it particularly well-suited for school-level tournaments, where detailed statistics are often unavailable. More broadly, these results underscore the potential of network-based methods to generate interpretable and stable assessments of relative team strength in constrained competition environments.

5.2. Contribution of the Work

This study contributes to the field of sports analytics by extending the application of the well-established PageRank algorithm beyond traditional benchmarking toward evaluating team strength in incomplete tournament environments. Rather than proposing methodological innovations, the research demonstrates how existing network-based techniques can be effectively adapted and validated in data-sparse, structurally constrained educational sports settings. Specifically, this study highlights PageRank's utility as a longitudinal assessment tool for high school baseball tournaments, addressing the unique analytical challenges posed by unbalanced match schedules and limited performance data.

The research establishes empirically grounded temporal parameters for youth sports contexts, identifying 3–4 years as the optimal historical window for PageRank-based predictions. This finding reflects the unique characteristics of high school athletics, including rapid player turnover and program evolution, where extending historical windows beyond four years introduces noise rather than signals due to structural changes and external disruptions.

These contributions distinguish this work from prior studies that primarily focus on inter-model comparisons. This study prioritizes algorithm interpretability and stakeholder acceptance—considerations that are crucial for real-world implementation in educational settings but often overlooked in comparative algorithmic studies.

5.3. Practical Implications

The application of PageRank in this study presents several practical implications for sports administrators and tournament organizers. First, the algorithm provides a low-cost and transparent method for ranking teams in the absence of detailed player-level statistics or complete round-robin results, thereby contributing to greater perceived fairness in youth-level competitions. This is particularly relevant for school-level tournaments, where budget and data infrastructure are often limited. Second, the approach can support more equitable seeding procedures and post-tournament evaluations by capturing the quality of victories rather than relying solely on their quantity. Finally, the identification of long-term strong programs through PageRank trends may aid in talent development and resource allocation across regions or institutions.

Additionally, the interpretability of PageRank rankings provides practical value for non-technical stakeholders such as coaches and school administrators. For coaches, the rankings can inform pre-match preparation by providing a data-based assessment of opponent strength, allowing for more strategic resource allocation, roster planning, and training focus. In high-stakes matchups, the rankings may also serve as a motivational tool.

For school administrators, these rankings provide a transparent and quantifiable measure of team performance that can support internal evaluations, justify athletic resource planning, and inform decisions related to coaching support or program development. The longitudinal use of such rankings can also help identify performance trends across years and guide institutional strategies for interscholastic sports development.

5.4. Limitations and Future Research

This study also has several limitations. The PageRank model used here is based solely on binary win–loss outcomes, without incorporating contextual factors such as the margin of victory, home-field advantage, or player-level variables. Furthermore, this study’s empirical scope is confined to a single tournament in Taiwan, which may limit the generalizability of its conclusions to other sports contexts or educational systems. Cross-validation against alternative sports or geographical contexts would strengthen external validity and broaden the applicability of these findings.

Additionally, this study did not conduct formal statistical robustness testing such as sensitivity analysis, bootstrap validation, or perturbation studies. While our findings demonstrate consistent performance across multiple seasons and temporal horizons, formal robustness evaluation remains an important direction for future research to establish PageRank’s reliability under data uncertainty in tournament settings.

Future studies may expand upon this work in several directions. Methodologically, incorporating weighted metrics such as score differentials or recency-weighted outcomes may enhance the sensitivity of PageRank in capturing real-time team dynamics. Time-decay variants or dynamic PageRank adaptations may also better accommodate performance evolution throughout a season. Additionally, comparative evaluations with alternative ranking models, including Elo, Colley, or Massey, would help clarify the strengths and trade-offs of different algorithms. Finally, applying these methods across different sports, competition levels, or international school systems can further test the scalability and versatility of network-based ranking approaches in sports analytics.

Author Contributions: Conceptualization, methodology, investigation, formal analysis, writing, and supervision, Y.-C.H.; data curation, W.-J.Z.; software, Y.-C.H. and W.-J.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the National Taiwan University of Sport (grant number 114DG00108).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data were obtained from the official website of the Black Panther Cup (<https://blackpanthercup.tw/>, accessed on 4 May 2025) and the Chinese Taipei Baseball Association (<http://www.ctba.org.tw/>, accessed on 4 May 2025). The cleaned and processed datasets used in this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Sziklai, B.R.; Biró, P.; Csató, L. The Efficacy of Tournament Designs. *Comput. Oper. Res.* **2022**, *144*, 105821. [[CrossRef](#)]
2. Irons, D.J.; Buckley, S.; Paulden, T. Developing an Improved Tennis Ranking System. *J. Quant. Anal. Sports* **2014**, *10*, 109–118. [[CrossRef](#)]
3. Dabadghao, S.S.; Vaziri, B. The Predictive Power of Popular Sports Ranking Methods in the NFL, NBA, and NHL. *Oper. Res. Int. J.* **2022**, *22*, 2767–2783. [[CrossRef](#)]
4. Lewis, M. *Moneyball: The Art of Winning an Unfair Game*; W. W. Norton & Company: New York, NY, USA, 2004; ISBN 978-0-393-06623-4.

5. Stefani, R.T. Survey of the Major World Sports Rating Systems. *J. Appl. Stat.* **1997**, *24*, 635–646. [[CrossRef](#)]
6. James, B. *The New Bill James Historical Baseball Abstract*; Simon and Schuster: New York, NY, USA, 2010.
7. Colley, W. *Colley's Bias Free College Football Ranking Method*; Princeton University: Princeton, NJ, USA, 2002.
8. Massey, K. *Statistical Models Applied to the Rating of Sports Teams*. Bachelor's Thesis, Bluefield College, Bluefield, VA, USA, 1997.
9. Chartier, T.; Harris, J.; Hutson, K.; Langville, A.; Martin, D.; Wessell, C. Reducing the Effects of Unequal Number of Games on Rankings. *IMAGE Bull. Int. Linear Algebra Soc.* **2014**, *52*, 15–23.
10. Park, J.; Yook, S.-H. Bayesian Inference of Natural Rankings in Incomplete Competition Networks. *Sci. Rep.* **2014**, *4*, 6212. [[CrossRef](#)]
11. Bozóki, S.; Csató, L.; Temesi, J. An Application of Incomplete Pairwise Comparison Matrices for Ranking Top Tennis Players. *Eur. J. Oper. Res.* **2016**, *248*, 211–218. [[CrossRef](#)]
12. Govan, A.Y.; Meyer, C.D.; Albright, R. Generalizing Google's PageRank to Rank National Football League Teams. In Proceedings of the SAS Global Forum, San Antonio, TX, USA, 16–19 March 2008; Volume 2008.
13. Moteqi, S.; Masuda, N. A Network-Based Dynamical Ranking System for Competitive Sports. *Sci. Rep.* **2012**, *2*, 904. [[CrossRef](#)]
14. Brown, S. A PageRank Model for Player Performance Assessment in Basketball, Soccer and Hockey. *arXiv* **2017**, arXiv:1704.00583. [[CrossRef](#)]
15. Elo, A.E. *The Rating of Chessplayers, Past and Present*; Arco Publishing: New York, NY, USA, 1978; ISBN 978-0-668-04721-0.
16. Glickman, M.E. The Glicko System. *Boston Univ.* **1995**, *16*, 9.
17. Bradley, R.A.; Terry, M.E. Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika* **1952**, *39*, 324–345. [[CrossRef](#)]
18. Keener, J.P. The Perron–Frobenius Theorem and the Ranking of Football Teams. *SIAM Rev.* **1993**, *35*, 80–93. [[CrossRef](#)]
19. Park, J.; Newman, M.E.J. A Network-Based Ranking System for US College Football. *J. Stat. Mech.* **2005**, *2005*, P10014. [[CrossRef](#)]
20. Stefani, R. The Methodology of Officially Recognized International Sports Rating Systems. *J. Quant. Anal. Sports* **2011**, *7*, 10. [[CrossRef](#)]
21. Barrow, D.; Drayer, I.; Elliott, P.; Gaut, G.; Osting, B. Ranking Rankings: An Empirical Comparison of the Predictive Power of Sports Ranking Methods. *J. Quant. Anal. Sports* **2013**, *9*, 187–202. [[CrossRef](#)]
22. Brin, S.; Page, L. The Anatomy of a Large-Scale Hypertextual Web Search Engine. *Comput. Netw. ISDN Syst.* **1998**, *30*, 107–117. [[CrossRef](#)]
23. Shi, J.; Tian, X.-Y. Learning to Rank Sports Teams on a Graph. *Appl. Sci.* **2020**, *10*, 5833. [[CrossRef](#)]
24. Yun, H.-J.; Park, J.-H.; Yoon, J.; Jeon, M. The Evaluation of the Team Performance of MLB Applying PageRank Algorithm. *Ann. Appl. Sport Sci.* **2022**, *10*, e968. [[CrossRef](#)]
25. Beggs, C.B.; Shepherd, S.J.; Emmonds, S.; Jones, B. A Novel Application of PageRank and User Preference Algorithms for Assessing the Relative Performance of Track Athletes in Competition. *PLOS ONE* **2017**, *12*, e0178458. [[CrossRef](#)]
26. Kim, E.; Jeon, M. Proposal for Implementation of a Ranking Model for Olympic Taekwondo Competitions Using PageRank. *Int. J. Perform. Anal. Sport* **2019**, *19*, 227–235. [[CrossRef](#)]
27. Oh, H.; Jeon, M.; Chin, S.; Lim, H. Application of PageRank Model for Olympic Women's Taekwondo Rankings: Comparison of PageRank and Accumulated Point Index System. *Ann. Appl. Sport Sci.* **2022**, *10*, e971. [[CrossRef](#)]
28. Zhou, Y.; Wang, R.; Zhang, Y.-C.; Zeng, A.; Medo, M. Improving PageRank Using Sports Results Modeling. *Knowl. Based Syst.* **2022**, *241*, 108168. [[CrossRef](#)]
29. Lee, S.; Kang, M.; Park, J.-H.; Yun, H.-J. Model Development for an Artificial Intelligence-Based NCAA Football Ranking System: Applying the PageRank Algorithm. *Meas. Phys. Educ. Exerc. Sci.* **2025**, 1–11. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.