Correlation Analysis of Ranking Methods via Transitive Triads in Random Tournament Graphs

Liu Bowen

Abstract. This study builds upon prior research, aiming to propose alternative ranking methods to improve fairness compared to simple win-count scoring. The key motivation arises from scenarios where a higher-ranked player (e.g., b) may accumulate wins against weaker opponents, while a lower-ranked player (e.g., a) achieves fewer but more significant victories against stronger competitors. In such cases, traditional win-count ranking may inaccurately reflect true skill, leading to potential unfairness.

In this paper, we analyze the correlation between two ranking methods in random tournament graphs, focusing on the role of transitive triads (or their absence, i.e., non-transitive triads). Our goal is to determine whether alternative scoring methods reduce dependence on raw win counts, thereby improving fairness. Through computational experiments, we investigate the structural properties of tournaments where the correlation between rankings is minimized, providing insights into the conditions under which alternative methods may offer more equitable outcomes.

This research aims to contribute empirical data for future studies assessing whether refined ranking approaches enhance fairness in competitive settings.

Introduction

This study builds upon prior research [1, 2] and aims to propose alternative ranking methods to improve fairness compared to simple win-count scoring. The core motivation stems from scenarios where a higher-ranked player (e.g., b) may accumulate wins against weaker opponents, while a lower-ranked player (e.g., a) achieves fewer victories but against stronger competitors. In such cases, traditional win-count ranking may fail to accurately reflect true skill, thereby leading to potential unfairness. In this paper, we analyze the correlation between two ranking methods in random tournament graphs, with a particular focus on the role of transitive triads (and their absence, i.e., non-transitive triads). Through computational experiments, we generate a large number of random tournament graphs to simulate diverse competitive environments. We then modify each triad within these graphs to create variants, and for each variant, we calculate ranking scores using two different ranking methods. Subsequently, we assess the correlation between these ranking scores using five algorithms: Spearman's rank correlation coefficient[3, 4], Kendall's tau coefficient[3, 4], Modified Kendall's tau, and three custom correlation metrics proposed by our research team. By examining the structural properties of these tournament graphs, we aim to determine whether alternative scoring methods reduce dependence on raw win counts, thereby improving fairness. We statistically evaluate the mean and variance of the correlation values obtained from these algorithms and identify the conditions under which lower correlation values emerge. Our goal is to provide empirical data to support future studies assessing whether refined ranking approaches can enhance fairness in competitive settings. Through the above analysis, we strive to provide valuable insights for the development of fairer ranking systems in competitive contexts.

1. Ranking Points Rules

After the game concludes, players' ranking points are calculated using two distinct methods:

1. **Standard Ranking (SR)**:
   - Players are sorted in descending order by their total points
   - The top-ranked player receives n ranking points
   - Each subsequent player receives one fewer ranking point ($n-1$, $n-2$, etc.)
   - Tied players receive the average of ranking points for their positions
     Example: If players ranked 2nd-4th tie, each receives $\frac{(n-1)+(n-2)+(n-3)}{3}$
2. **Alternative Ranking (AR)**:
   - Let P be the set of all n players
   - For every non-empty subset $S \subseteq P$ where $|S| \geq 2$:
     (a) Conduct virtual tournaments within S using existing duel results
     (b) Award standard ranking points to members of S based on their subset performance
     (c) These subset ranking points are temporary and denoted as $r_S(p)$ for player $p \in S$
   - Each player's aggregate score is computed as:

$$R(p) = \sum_{\substack{S \subseteq P \\ p \in S \\ |S| \geq 2}} r_S(p)$$

   - Final ranking points are determined by applying the SR method to the $R(p)$ scores

2. Research Plan

To achieve our research objectives, we randomly generated 1,000,000 Tournaments through a program and calculated the ranking points for each graph using these methods. By examining the structural properties of these tournament graphs, we aim to determine whether alternative scoring methods reduce dependence on raw win counts, thereby improving fairness. We statistically evaluate the mean and variance of the correlation values obtained from these algorithms and identify the conditions under which lower correlation values emerge. Our goal is to provide empirical data to support future studies assessing whether refined ranking approaches can enhance fairness in competitive settings. Through the above analysis, we strive to provide valuable insights for the development of fairer ranking systems in competitive contexts.

Now, we consider the role of transitive triads in this invariant. There are four ways to change a transitive triad:

1. Change the direction of the first edge, keeping other edges unchanged.
2. Change the direction of the second edge, keeping other edges unchanged.
3. Change the direction of the third edge, keeping other edges unchanged.
4. Change the direction of all edges.

Here, changing the direction of edges means altering the win-loss results of the participants' duels. The uniqueness of the fourth change method lies in the fact that the total number of wins for each participant remains unchanged (i.e., the ranking points calculated using the first method do not change), but the ranking points calculated using the second method may change. Specifically, during the calculation of the second ranking method, there exist sub-tournaments that only contain the two participants involved in the fourth change. In these sub-tournaments, the total number of wins for these participants changes, which alters their ranking points calculated by the first method. This change may not be compensated by other sub-tournaments, as each ranking point and its corresponding number of wins are only correlated within the current sub-tournament graph, not equal. Therefore, this method may impact the overall ranking points.

To statistically evaluate the correlation between different ranking methods, we employed Spearman's rank correlation coefficient, Kendall's tau coefficient, Modified Kendall's tau, and three custom correlation metrics proposed by our research team. We statistically evaluate the mean and variance of the correlation values obtained from these algorithms and identify the conditions under which lower correlation values emerge.

Experimental Data

The following tables present the statistical data for each correlation method used in our experiments:

Correlation Method	Mean	Variance	Standard Deviation
Spearman's rank correlation coefficient	0.965954	0.00111474	0.0333877
Kendall's tau coefficient	0.965954	0.00111474	0.0333877
Modified Kendall's tau	1	0	0
Custom Metric 1	0.843738	0.00758343	0.0870829
Custom Metric 2	2.93272	0.103837	0.322237

Table 1. Statistical Data for Correlation Methods

Conclusion

This study has explored alternative ranking methods to address the limitations of traditional win-count scoring in competitive settings. By generating a large number of random tournament graphs and systematically modifying their transitive triads, we evaluated the correlation between two ranking methods using various statistical metrics. Our results indicate that certain transformations, especially those involving the reversal of all edges in a triad, can impact ranking outcomes. However, further research is needed to fully understand the implications of these findings. Overall, our findings provide valuable directions for future exploration.

References

[1] Bowen Liu, *Comparison of two methods for calculating ranking points using transitive triads*, **9**(4), 309–320 (2024), doi:10.22061/jdma.2024.11215.1093. Available at `https://jdma.sru.ac.ir/article_2222.html`.

[2] Melnikov, B. and Liu, B. (2024) On an Invariant of Tournament Digraphs. Journal of Applied Mathematics and Physics, 12, 2711-2722. doi: 10.4236/jamp.2024.127162.

[3] M. Lagutin, *Visual Mathematical Statistics*, Binom: Moscow, Russia, 2012, 472 p. (In Russian).

[4] L. Wasserman, *All of Statistics: A Concise Course in Statistical Inference*, Springer Science & Business Media: Berlin, Germany, 2013, 442 p.

Liu Bowen
Department of Mathematics, Shenzhen MSU-BIT University, Shenzhen, China.
e-mail: `liubowenmbu@163.com`