Hing-King-Linden Score: A robust method of assessing foosball matchups independent of tournament structure

Philip Linden, Adam King and Brandon Hing

Abstract— **The results of randomly seeded single-elimination tournaments does not reflect a player or team's relative rank, or foosball performance score. Existing ranking such as Elo Ranking or Swiss systems depend on a large number of matches to be played to produce accurate results. We present a method of determining rank using the relative performance of a player's opponent and the points scored on that opponent during a foosball match. We demonstrate that the Hing-King-Linden Score, or HKL, is a robust method of predicting foosball matchups between two players or teams, and a reliable way to seed teams and tournaments fairly based on only the number of matches played and tracking points scored against each opponent. We show the utility of HKL using the real results of a randomly seeded single-elimination tournament.**

I. INTRODUCTION

Championship tournaments are a tried-and-true organization for competitions in sports today, and foosball is no exception. However, both single- and double-elimination tournament results are predicated upon fair seeding of the matchups at the start of the tournament. Some existing ranking systems used to seed tournaments are Elo Ranking[], used for the FIFA World Cup[], or the Swiss-system[], commonly used in chess tournaments[]. A fair score in both Elo and Swiss systems requires that each member of the tournament plays every other in round-robin fashion, where $\frac{n}{2}(n-1)$ matches must be played before a rank is determined. The difference in these rankings is in how scores incorporate margin of victory and win/loss rates. Additionally, existing rating systems assume that results are binary (win or loss) or uncapped scores (higher score wins). In foosball, wins are determined by the first team to score a set number of points.

In a recent tournament, the competition organizers used a single-elimination tournament between 16 randomly seeded players to seed an 8-team tournament of evenly matched 2-person teams. In this event structure, we must determine relative skill of 16 players when the conditions for Elo and Swiss rankings systems are not met—not only have players not played at least one match against one another, but some players played more matches than others. In some cases, players eliminated in the first round scored more points and lost by a smaller margin of victory against the overall winner of the tournament. Thus it is clear that a robust ranking system is needed to assemble fair teams based on match results using the relative performance of a player's opponent to determine that player's rank.

In this paper we propose the Hing-King-Linden Score, or HKL, which weighs the cumulative number of points scored against each opponent and the relative performance of that opponent. Critically, HKL only requires one randomly

seeded single-elimination tournament to determine an accurate score that reflects relative performance.

The primary goal of the HKL is to rank contestants fairly given the following principles:

1) Given two players who lose a match in the same round of the tournament by the same margin of victory, the player whose opponent progresses through more subsequent rounds should have a higher HKL.
2) The winner of a randomly-seeded single-elimination tournament does not necessarily have the highest rank.
3) Two players with similar HKL compete with similar performance.

II. DEFINITION

As matches are played, points scored are tracked individually for each player per opponent. Points scored per opponent are cumulative, but in a single-elimination tournament no pair of contestants appears in the bracket more than once.

v scored on >	Abhra	Adam	Ben	Bethany	Brandon	Caleb	Cruz	Jeff	Kyle	Lisa	Matt	Natasha	Phil	Sean	Steve	Tyler	Total points scored
Abhra		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Adam	0		0	0	0	0	0	0	0	0	0	0	0	4	0	0	4
Ben	0	0		0	5	0	3	0	0	0	0	0	0	0	5	0	13
Bethany	0	0	0		0	4	0	0	0	0	0	0	0	0	0	0	4
Brandon	0	0	1	0		0	0	0	0	0	0	0	0	5	0	0	6
Caleb	0	0	0	5	0		5	0	0	0	0	0	0	5	0	5	20
Cruz	0	0	5	0	0	2		0	5	0	5	0	0	0	0	0	17
Jeff	0	0	0	0	0	0	0		0	0	0	2	0	0	0	0	2
Kyle	0	0	0	0	0	0	1	0		0	0	0	0	0	0	0	1
Lisa	0	0	0	0	0	0	0	0	0		4	0	0	0	0	0	4
Matt	0	0	0	0	0	0	2	0	0	5		0	0	0	0	0	7
Natasha	0	0	0	0	0	0	0	5	0	0	0		0	0	0	2	7
Phil	0	0	0	0	3	0	0	0	0	0	0	0		0	0	0	3
Sean	0	5	0	0	0	1	0	0	0	0	0	0	0		0	0	6
Steve	0	0	1	0	0	0	0	0	0	0	0	0	0	0		0	1
Tyler	5	0	0	0	0	1	0	0	0	0	0	5	0	0	0		11

Fig. 1. Points scored are tracked per opponent for each player.

Since some players progress further than others in the bracket and thus have more scoring opportunities in every round, it is necessary to normalize the total points scored per opponent, p, by a player's progression in the tournament, r, where r is the number of matches that a player has completed.

$$\hat{p} = p/r \tag{1}$$

A player's overall performance, k, is then calculated as the root-sum-square of the normalized total points scored and progression in the tournament. By using the root-sum-square of progression and normalized scores, we estimate how consistently a player performs in the tournament. A

Normalized score	Abhra	Adam	Ben	Bethany	Brandon	Caleb	Cruz	Jeff	Kyle	Lisa	Matt	Natasha	Phil	Sean	Steve	Tyler	Adj. total score
Abhra		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Adam	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0	0.0	0.0	4.0
Ben	0.0	0.0		0.0	1.7	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.7	0.0	4.3
Bethany	0.0	0.0	0.0		0.0	4.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	4.0
Brandon	0.0	0.0	0.5	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.5	0.0	0.0	0.0	3.0
Caleb	0.0	0.0	0.0	1.3	0.0		1.3	0.0	0.0	0.0	0.0	1.3	0.0	1.3	0.0	1.3	5.0
Cruz	0.0	0.0	1.3	0.0	0.0	0.5		0.0	1.3	0.0	1.3	0.0	0.0	0.0	0.0	0.0	4.3
Jeff	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0	0.0	0.0	2.0	0.0	0.0	0.0	0.0	2.0
Kyle	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0		0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0
Lisa	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		4.0	0.0	0.0	0.0	0.0	0.0	4.0
Matt	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	2.5		0.0	0.0	0.0	0.0	0.0	3.5
Natasha	0.0	0.0	0.0	0.0	0.0	0.0	0.0	2.5	0.0	0.0	0.0		0.0	0.0	0.0	1.0	3.5
Phil	0.0	0.0	0.0	0.0	3.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0	0.0	0.0	3.0
Sean	0.0	2.5	0.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0	0.0	3.0
Steve	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0		0.0	1.0
Tyler	1.7	0.0	0.0	0.0	0.0	0.3	0.0	0.0	0.0	0.0	0.0	1.7	0.0	0.0	0.0		3.7

Fig. 2. Scored points are normalized by progression in the tournament.

Matchup difficulty	Abhra	Adam	Ben	Bethany	Brandon	Caleb	Cruz	Jeff	Kyle	Lisa	Matt	Natasha	Phil	Sean	Steve	Tyler
Abhra	1.00	4.12	5.27	4.12	3.61	6.40	5.84	2.24	1.41	4.12	4.03	4.03	3.16	3.61	1.41	4.74
Adam	0.24	1.00	1.28	1.00	0.87	1.55	1.42	0.54	0.34	1.00	0.98	0.98	0.77	0.87	0.34	1.15
Ben	0.19	0.78	1.00	0.78	0.68	1.21	1.11	0.42	0.27	0.78	0.76	0.76	0.60	0.68	0.27	0.90
Bethany	0.24	1.00	1.28	1.00	0.87	1.55	1.42	0.54	0.34	1.00	0.98	0.98	0.77	0.87	0.34	1.15
Brandon	0.28	1.14	1.46	1.14	1.00	1.78	1.62	0.62	0.39	1.14	1.12	1.12	0.88	1.00	0.39	1.31
Caleb	0.16	0.64	0.82	0.64	0.56	1.00	0.91	0.35	0.22	0.64	0.63	0.63	0.49	0.56	0.22	0.74
Cruz	0.17	0.71	0.90	0.71	0.62	1.10	1.00	0.38	0.24	0.71	0.69	0.69	0.54	0.62	0.24	0.81
Jeff	0.45	1.84	2.36	1.84	1.61	2.86	2.61	1.00	0.63	1.84	1.80	1.80	1.41	1.61	0.63	2.12
Kyle	0.71	2.92	3.73	2.92	2.55	4.53	4.13	1.58	1.00	2.92	2.85	2.85	2.24	2.55	1.00	3.35
Lisa	0.24	1.00	1.28	1.00	0.87	1.55	1.42	0.54	0.34	1.00	0.98	0.98	0.77	0.87	0.34	1.15
Matt	0.25	1.02	1.31	1.02	0.89	1.59	1.45	0.55	0.35	1.02	1.00	1.00	0.78	0.89	0.35	1.18
Natasha	0.25	1.02	1.31	1.02	0.89	1.59	1.45	0.55	0.35	1.02	1.00	1.00	0.78	0.89	0.35	1.18
Phil	0.32	1.30	1.67	1.30	1.14	2.02	1.85	0.71	0.45	1.30	1.27	1.27	1.00	1.14	0.45	1.50
Sean	0.28	1.14	1.46	1.14	1.00	1.78	1.62	0.62	0.39	1.14	1.12	1.12	0.88	1.00	0.39	1.31
Steve	0.71	2.92	3.73	2.92	2.55	4.53	4.13	1.58	1.00	2.92	2.85	2.85	2.24	2.55	1.00	3.35
Tyler	0.21	0.87	1.11	0.87	0.76	1.35	1.23	0.47	0.30	0.87	0.85	0.85	0.67	0.76	0.30	1.00

Fig. 3. The ratio of d for any matchup indicates the relative difficulty for one player against another.

high performing player is likely to win many rounds of the tournament by large margins while a weaker player might win some matches but by small margins. Conversely, a player that scores many points in just a few matches is more likely to be a strong player.

$$k = \sqrt{\hat{p}^2 + r^2} \qquad (2)$$

To provide a basis for comparison between player performance scores, we linearly scale k to an arbitrary fixed range, such as $n = (0, 10)$, to compose a player's Power Rating, H. Since this scale is relative, scaling depends on the maximum and minumum k in the set of all players in the tournament.

$$K = n_1 + \left(\frac{n_2 - n_1}{\max(k) - \min(k)} \right) (k - \min(k)) \qquad (3)$$

Since k is independent of a player's opponent, we can estimate the relative difficulty of any matchup between two players, d, with a simple ratio. A difficulty of $d < 1$ indicates an "easy" match while $d > 1$ indicates a "difficult" match from the perspective of any player against any opponent. A perfectly even matchup results in $d = 1$, which makes sense if considering the most perfect matchup—a player versus themselves—is unity difficulty. For reporting, we may scale difficulty using Equation 3, replacing k with d.

$$d = \frac{k_{player}}{k_{opponent}} \qquad (4)$$

Next, we weigh the number of points scored per opponent by the matchup difficulty against that opponent. In other words, a point scored in a difficult match is "worth" more than a point scored in an easy match, regardless of which round of the tournament the point was scored. We estimate a player's cumulative performance, s, by summing the total adjusted points, $\bar{p}$, scored in all rounds of the tournament. Again, we linearly scale s to arbitrary range n to ground it in a relative basis, and map it to the same scale as K. The scaled performance metric, S, is called a player's Seed Rating.

$$\bar{p} = pd \qquad (5)$$

$$s = \sum \bar{p} \qquad (6)$$

$$S = n_1 + \left(\frac{n_2 - n_1}{\max(s) - \min(s)} \right) (s - \min(s)) \qquad (7)$$

Finally, we assert that Power Rating, K, and Seed Rating, S, are equally justified as indicators of a player's performance. We see that Power Rating estimates performance over many matches while Seed Rating accounts for performance in individual matches. Thus we arrive at the HKL, H, defined as the average between the Power and Seed Ratings. Since K and S are scaled to the same fixed range, H is also scaled to that range.

$$H = (K + S)/2 \qquad (8)$$

Player	Parameters			S	Rankings	H
	r	$\bar{p}$	k	S	K	H
Abhra	1	0.00	1.00	0.00	0.00	0.00
Adam	1	4.00	4.12	2.45	5.78	4.11
Ben	3	4.33	5.27	5.66	7.90	6.78
Bethany	1	4.00	4.12	4.35	5.78	5.06
Brandon	2	3.00	3.61	4.09	4.82	4.46
Caleb	4	5.00	6.40	10.00	10.00	10.00
Cruz	4	4.25	5.84	7.96	8.95	8.45
Jeff	1	2.00	2.24	2.52	2.29	2.41
Kyle	1	1.00	1.41	2.89	0.77	1.83
Lisa	1	4.00	4.12	2.74	5.78	4.26
Matt	2	3.50	4.03	5.60	5.61	5.61
Natasha	2	3.50	4.03	3.59	5.61	4.60
Phil	1	3.00	3.16	2.39	4.00	3.20
Sean	2	3.00	3.61	5.24	4.82	5.03
Steve	1	1.00	1.41	2.61	0.77	1.69
Tyler	3	3.67	4.74	4.66	6.92	5.79

TABLE I

EXAMPLE HKL AND INTERMEDIATE PARAMETERS

III. FEATURES

Since HKL is relative between all members of a tournament and only depends on scores earned during matches, three major advantages emerge from this new ranking method:

1) HKL of all contestants is defined as soon as one point is scored in the entire tournament.
2) HKL can be updated during a match that has not yet completed.
3) HKL is independent of which matches on the tournament schedule are played, and independent of the order that they are played.

Another feature of the HKL that it is purely additive in nature. Even if a strong player loses a match to a opponent with a much lower rank, the stronger player still earns points. In other words, all parameters are greater than or equal to zero. This means that the only way a player's HKL can decrease is if the minimum or maximum of all k or all s change.

A. Example Scenario: No decrease in HKL after a match

Two players, A and B, play a match with a matchup diffulty $d = 1$. At the start of the match, their performance parameters, k and s, are greater than the performance parameters of player C, but less than the performance parameters of player D, who has the highest HKL out of all players in the tournament. The final score is $5 - 0$ and player A wins the match. The number of matches completed, r, for both players increases by 1. Even though player B did not score any points, it is not possible for their k or s to decrease, and the performance parameters of both players is still higher than player C. Player A has scored enough adjusted points, $\bar{p}$, to tie player D in s but has played more matches than player D and still has a lower k score. In this example, the minimum and maximum s and k out of all players in the tournament has not changed. Therefore both player A and player B gain in Seed Rating, S, and Power Rating, K, but no other players in the tournament see a decrease in HKL.

IV. CASE STUDY: PREDICTION ACCURACY OVER A SINGLE-ELINIMATION TOURNAMENT

Like other ranking systems, HKL accuracy increases with the number of matches played. In this section, we demonstrate how HKL predicts match results and how these predictions respond to match results using real data collected during a single elimination tournament.