Monte-Carlo Simulation of Chess Tournament Classification Systems

Tanja Van Hecke

University College Ghent, Department of Applied Engineering Sciences, V. Vaerwyckweg 1, B-9000 Ghent, Belgium

Abstract   What makes you win a chess tournament? You should be talented, but the choice of opponents and the applied ranking system will influence your final ranking as well. This article compares different ranking systems (such as the Swiss system and the elimination system) and makes simulations to judge which system is fair. It will lead to a compromise between a qualitative representation of the level of the participants and the length of the tournament or the required number of duels. We constructed an error measurement to evaluate the methods. Also by means of the 95% confidence interval, it was possible to detect the superior qualities of the Swiss system.

Keywords   Monte-Carlo Simulation, Ranking System, Estimation, Error Function

1. Swiss System Versus Elimination System

Many games (chess, bridge, tennis, ...) confront two players. When tournaments take place, many players have to face each other. The way of pairing should lead to a correct classification. A "Round Robin" is a format of chess tournaments where each opponent plays all of the other opponents. This is the best way of determining playing strength; however, the number of rounds needed are prohibitive for a large number of entrants. For example, for 16 players, there would be 15 rounds using the Round Robin format.

Ranking participants at tournaments has been described and discussed from various perspectives [6][7][8][9]. We will develop an evaluation function for ranking methods based on Monte-Carlo simulations. To obtain a representative classification, there exist two main systems: the elimination system and the Swiss system[2][5]. The basic rule for the elimination system is once you lose a round, you fall out and will no longer play. This reduces the number of confrontations between players and is often used in tennis tournaments with infrastructural limitations. The pairing occurs at random at all rounds of the game. Chess tournaments require much less infrastructure. Here the Swiss tournament system is chosen more often. It was invented by J. Muller and first used in a chess tournament at Zurich (Switzerland) in 1895 (hence "Swiss" system). We assume that players are paired ad random for the starting round. There exist systems where this assumption is replaced by an initial classification based on results of the tournament of the previous year, e.g. the McMahon system. As only the winning players continue, the last round is the most spectacular one where the two best players duel. In the Swiss system the players are classified by the number of games they have won so far. When the number of players is a power of 2, i.e. $m = 2^k$, it is possible to select pairs of players who both won the same number of games, which makes this case the most preferable. We assume one may not play the same player twice within the same tournament.

2. Tie Breaks

The final classification in the Swiss system is based firstly on the total number of games won by the player. But this doesn't make the players unique and tie breaks will occur. To make a further ranking within the sets S_j of players who won the same number j of games, a refinement of the Swiss system is required. Two important tie break methods are *the Solkoff method* and *the cumulative method*[3]. The Solkoff method is based on the strength of opponents a player has beaten and is quantified by a weight quantity w_i. This weight quantity w_i of player i can be defined as the sum of the number of games won by the players beaten by player i. Within a set S_j players with higher weights get smaller ranking numbers. For those players with equal weights, the mean of the concerned weight is attributed to each of them. For example: when the ranking numbers 10 and 11 are to be attributed to two players within the same set with equal weights, they are both ranked by number 10.5.

The cumulative method is based on the sum sc_i of the cumulative (running) scores for each round of player i. So if player i won his first 3 games, lost the fourth game, and won the fifth, his cumulative score is: 1 2 3 3 4. This makes $sc_i = 1 + 2 + 3 + 3 + 4 = 13$, which would be his

* Corresponding author:
Tanja.VanHecke@hogent.be (Tanja Van Hecke)

cumulative tiebreak. If player i lost his first game, and won the last four, he would have a cumulative tiebreak as follows: $sc_i = 0 + 1 + 2 + 3 + 4 = 10$, tiebreak points. The reasoning behind this method is based on the Swiss system of playing an opponent with the same score as oneself. The assumption is that if you win early, you're playing tougher opponents (opponents who also won early and probably finished higher). If you lost in the early rounds, you played weaker opponents (who also lost early and probably didn't finish as high). The higher sc_i, the better player i.

3. The Game Process

3.1. Simulation Condition

The best way to make a comparative analysis of several ranking methods, is in case the number of players is a power of 2. For example, we consider 16 participants numbered from 1 to 16 with the assumption that the smaller the number, the better the player. This means that when player i and player j are confronted, player i will win if $i < j$. As the goal of a tournament is to create a classification of players that represents their level, we expect as outcome of the tournament: player i ends up at ranking i.

3.2. Indicating a Winner

The number of active players in each round is shown in Table 1 for both systems when the tournament starts with 16 players. As $m = 2^4$, a winner can be indicated after four rounds.

Table 1.   number of players during the different rounds

method	Round 1	round 2	round 3	round 4
elimination	16	8	4	2
Swiss	16	16	16	16

After i rounds, $m\, C_i^j / 2^i$ players were able to win j games with the Swiss system, as is shown in Table 3 where the cardinality of the sets S_i is given at the different stages of the tournament. Here S_i is the set of players who have won i games so far.

Table 2 gives analogous cardinalities with the elimination system.

Table 2.   Cardinality of S_i ($i \in \{0, 1, 2, 3, 4\}$) after several rounds with the elimination system

Set	round 1	round 2	round 3	round 4
S_0	8	8	8	8
S_1	8	4	4	4
S_2	0	4	2	2
S_3	0	0	2	1
S_4	0	0	0	1

To compare and evaluate the two basic systems of tournament classification, we simulate the Swiss method

without secondary ranking within the sets S_i. They all receive the same mean ranking of the set they are in (see Table 4 based on the cardinalities of Table 2 and 3 to determine the mean values). For example, for the Swiss system the ranking 8.5 is assigned to all elements of set S_2 after 4 rounds as this set contains 6 elements which should receive the rankings 6, 7, 8, 9, 10 and 11. As no secondary ranking within the set is taken into account here, all elements get the mean ranking 8.5.

Table 3.   Cardinality of S_i ($i \in \{0, 1, 2, 3, 4\}$) after several rounds with the Swiss system

Set	round 1	round 2	round 3	round 4
S_0	8	4	2	1
S_1	8	8	6	4
S_2	0	4	6	6
S_3	0	0	2	4
S_4	0	0	0	1

Table 4.   Awarded ranking for elements of S_i ($i \in \{0, 1, 2, 3, 4\}$) at the end of the tournament

System	S_4	S_3	S_2	S_1	S_0
Elimination	1	2	3.5	6.5	12.5
Swiss	1	3.5	8.5	13.5	16

4. Numerical Results

4.1. Elimination Method

We discuss an example with 16 players in detail where a random generator determines the initial pairing:

(9, 8) (13, 5) (10, 1) (4, 16) (2, 3) (14, 6) (11, 12) (7, 15).

Applying the elimination method to this initial situation leads to

$S_0 = \{9, 13, 16, 10, 3, 14, 12, 15\}$
$S_1 = \{8, 5, 1, 4, 2, 6, 11, 7\}$

After randomization in S_1 the following pairing is proposed:

(5, 6), (4, 2), (11, 8), (1, 7)

After two rounds, the players are partitioned into the sets

$S_0 = \{9, 13, 16, 10, 3, 14, 12, 15\}$
$S_1 = \{6, 4, 11, 7\}$
$S_2 = \{5, 2, 8, 1\}$

After randomization the following pairing is proposed:

(2, 8) (1, 5)

After three rounds, the players are partitioned into the sets

$S_0 = \{9, 13, 16, 10, 3, 14, 12, 15\}$
$S_1 = \{6, 4, 11, 7\}$
$S_2 = \{5, 8\}$
$S_3 = \{2, 1\}$

A final fourth round will indicate player 1 as the winner. Remark that with the elimination method the major objective is indicating the winner. The ranking of the other participants

is of minor importance. As no supplementary effort is made to determine this complete ranking, there is a gain of number of duels, which is paid by an imprecise ranking of the participants other than the winner. In the example above, player 3 for example ends up in the set of lowest level players. No distinction can be made between him and player 16 and others.

4.2. Swiss System

We start by pairing the 16 players at random as in the previous section with the elimination method. The first round with the Swiss system gives the same sets S_0 and S_1 as in the previous section. After randomization in the two sets the following pairing is proposed:

(3, 9) (10, 15) (14, 16) (12, 13) (5, 6) (4, 2) (11, 8) (1, 7).

Remark that the pairing is made internally within the sets S_0 and S_1, in order to meet the principle to make duels of players of comparative level.

After two rounds, the players are partitioned into the sets
S_0= {9, 16, 15, 13}
S_1 = {3, 10, 14, 12, 6, 4, 11, 7}
S_2 = {5, 2, 8, 1}

After randomization the following pairing is proposed:
(13, 9) (16, 15) (14, 11) (7, 3) (4, 10) (6, 12) (2, 8) (1, 5).

After two rounds, the players are partitioned into the sets
S_0= {13, 16}
S_1 = {9, 15, 14, 7, 10, 12}
S_2 = {11, 3, 4, 6, 8, 5}
S_3 = {2, 1}

After a last randomization of the sets, the last round with pairs

(13, 16) (7, 9) (15, 14) (10, 12) (3, 6) (11, 4) (8, 5) (1, 2)

will determine the final ranking given in Table 5.

Table 5. Ranking of the different players with the Swiss system (basic and refined version with Solkoff and cumulative tie break methods)

Player i	r_i	w_i	r_i^s	sc_i	r_i^c
1	1	10	1	10	1
2	3.5	8	2	9	2
3	3.5	5	4.5	6	5
4	3.5	4	3	7	4
5	3.5	5	4.5	8	3
6	8.5	3	7	6	7.5
7	8.5	2	9.5	5	9
8	8.5	3	7	7	6
9	13.5	1	12.5	2	13.5
10	8.5	2	9.5	4	10.5
11	8.5	3	7	6	7.5
12	13.5	1	12.5	3	12
13	13.5	0	14.5	1	15
14	8.5	1	11	4	10.5
15	13.5	0	14.5	2	13.5
16	16	0	16	0	16

This table also mentions the refined ranking r_i^s and r_i^c heading tie breaks within the sets S_i ($i \in \{0, 1, 2, 3\}$) with respectively the Solkoff and the cumulative tie break methods. The final ranking with both the Solkoff and the cumulative method is not perfectly correct. For example in case of the cumulative method, player 3 receives ranking 5 and player 5 receives 3. They both lost one duel, but at different moments (for player 3 is was in an earlier round than for player 5). In case of the Solkoff method their ranking is equal as their opponents were of the same level.

In order to evaluate the different tournament classification systems, we need to measure the error of the created ranking. Therefore we consider the error

$$e = \sum_{i=1}^{m} (i - r_i)^2$$

where r_i is the ranking for player i.

When no secondary ranking is applied within the sets, the error e takes the value 77.5 for the Swiss system. It can be reduced to

$$\sum_{i=1}^{16} (i - r_i^s)^2 = 52$$

for the Solkoff method and to

$$\sum_{i=1}^{16} (i - r_i^c)^2 = 69.5$$

for the cumulative method. As this is just one sample, a better comparison of the methods will arise with a repetition of such a sample. Therefore the Monte Carlo simulation method is appropriate for our analysis.

4.3. Monte Carlo Simulation of a Chess Tournament

The Monte Carlo simulation[4],[1] technique uses multiple trial runs to discover statistical characteristic features. We apply this method to the tournament case and make several runs of a tournament starting from a random pairing of the players.

We consider an error e_j associated with run j defined as

$$e_j = \sum_{i=1}^{m} (i - r_i^j)^2,$$

where r_i^j is the ranking for player i in run number j of the simulation. It represents the summed square of the deviation of the expected ranking and the obtained value by each simulation run of the tournament with m players. The mean value $\bar{e}$ after n runs can be obtained by

$$\bar{e} = \frac{1}{n} \sum_{j=1}^{n} \sum_{i=1}^{m} (i - r_i^j)^2$$

and represents a global error of a ranking method.

Figure 1 illustrates the error associated with the elimination method as ranking method. Horizontally the number of the player can be found. Vertically the mean ranking $\bar{r_i}$ of a player i after $n = 200$ runs in the simulation of the tournament, can be found, where $\bar{r_i}$ is defined as

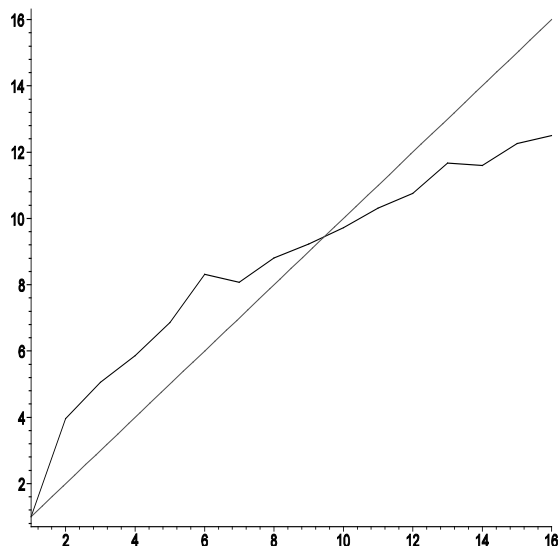$$\overline{r_i} = \frac{1}{n} \sum_{j=1}^{n} r_i^j.$$

Figure 1. The mean score after 200 runs with the elimination method versus expected value for the different players

The expected value is added by the straight line, as we made the assumption that the smaller the number of the player, the better the player. Figure 2 shows the same variables but here the ranking is made with the Swiss system.
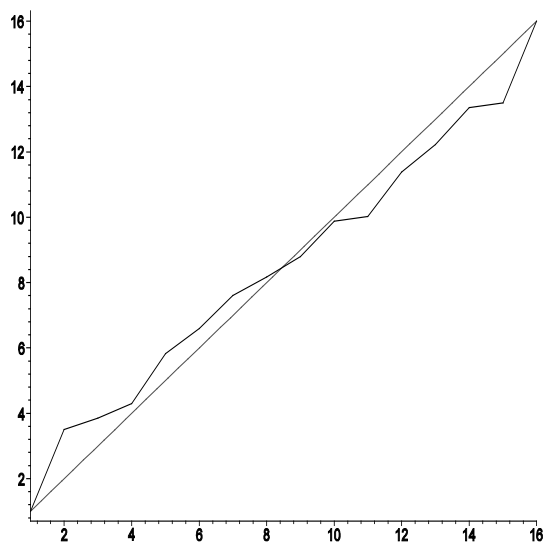

Figure 2. The mean score after 200 runs with the Swiss system versus expected value for the different players

Visually it is already clear that the approximation of the first angle bisector is better in case of the Swiss method compared to the elimination method. This is confirmed by the numbers for the mean error $\bar{e}$. Further statistics of e including its 95% confidence interval can be compared by means of Table 6.

Table 6. Statistics of the error e

System	$\bar{e}$	s_e	95% C.I. of e
Elimination	184.225	936.50	[54.432, 314.018]
Swiss	59.70	297.04	[18.533,100.867]

5. Conclusions

We were able to develop an error function associated with the Monte-Carlo simulation of tournament classification systems. This quantification revealed the superior classification feature of the Swiss system compared to the elimination system when ranking players at a chess tournament. The Swiss method requires the same number of rounds, but a larger number of duels. A ranking which perfectly represents the quality of the players is the Round Robin, but it has the disadvantage that it requires $n - 1$ rounds when n players are involved, which is not practicable. A more refined method with secondary ranking such as the Solkoff and cumulative method were compared in favor of the Solkoff method. Considering the strength of the opponents makes it possible to improve the ranking of the players.

REFERENCES

[1]   C.P. Robert, G. Casella, Monte Carlo Statistical Methods, 2nd ed., New York: Springer, 2010.

[2]   T. Just, D. Burg, Chess Federation's Official Rules of Chess, McKay, 2003.

[3]   www.winbeam.com/~chess/tiebreaks.html

[4]   N. Metropolis, S. Ulam, "The Monte Carlo method", Journal of Amer. Stat. Assoc., Vol. 44, pp.335—341, 1949.

[5]   A.W. Root, J.D. McNeil, "Children and chess: a guide for educators", London: Teacher Ideas Press, pp. 27—28, 2006.

[6]   J. Laslier, Tournament Solutions and Majority Voting, Springer-Verlag, Berlin, 1997.

[7]   A. Rubinstein, "Ranking the Participants in a Tournament," SIAM Journal on Applied Mathematics, Vol. 38, 108–111, 1980.

[8]   G. Slutzki, O. Volij, "Ranking Participants in Generalized Tournaments", International Journal of Game Theory, Vol. 33, 255–270, 2005.

[9]   M. Brozos-Vazquez, M.A. Campo-Cabana, J.C. Dıaz-Ramos, J. Gonzalez-Dıaz, "Ranking Participants in Tournaments by means of Rating Functions", Journal of Mathematical Economics, Vol. 44, 1246—1256, 2008.