Random Swiss Pairing System

Takahisa Kawai[1] and Hiroyuki Iida[1,2]
1 Department of Computer Science, Shizuoka University
2 PRESTO, Japan Science and Technology Agency
E-mail: {gs2016, iida}@cs.inf.shizuoka.ac.jp

Abstract

In every sport, pairing systems such as Swiss System are used to select the best player or the k-best players with the strict ordering among a set of competing players or teams. This paper examines the simulation-based outcome of several well-known pairing systems such as Round Robin, Swiss System and Random Swiss pairing. We then seek designs that are optimally efficient in the sense that the pairing system of limited rounds becomes closer to the outcome of Round Robin. Our models for the simulation reflect the fact that the better player in games between two players or teams does not always win, but it depends on their strength. The experiments performed show that Random Swiss pairing outperforms Swiss System.

1. Introduction

In many tournaments for games such as chess and sporting events throughout the world, various pairing systems are used to select suitable partners to facilitate accurate ranking, while determining the winner. In most cases Round Robin cannot be used, hence the assignment of partners highly influences the outcome of the tournament of limited rounds to play. Although the pairing is important, there has been little research addressing these issues statistically.

We then seek designs that are optimally efficient in the sense that the pairing system of limited rounds becomes closer to the outcome of Round Robin. In this paper several well-known pairing systems such as Round Robin, Swiss pairing and Random Swiss pairing are compared by simulation experiments.

2. Method

2.1. Pairing systems

We describe the basic idea of several pairing systems that we compare in this paper.

Round Robin

In the 1800's the format of chess tournaments was often Round Robin where each played all of the other entrants. This must be the best way to determine the best player or the k-best players with the strict ordering among a set of competing players or teams. However, the number of rounds needed is prohibitive for a large number of entrants.

Random Pairing

In Random pairing, the players to play each other will be determined just at random. However, two players shall not meet more than once.

Swiss System

It is invented by J. Muller and first used in a chess tournament at Zurich, Switzerland in 1895 (hence Swiss System)[1]. Swiss System is a system invented to compensate for the shortcoming. Its principle is to pair players with the nearest possible results in a tournament without playing the same opponent twice. Players must have played the same number of matches in order for them to play one another. The important features of Swiss System are listed below [1][2].

- In Swiss System, after the first round players are placed in groups according to their scores (winners in the group 1, those who drew go in the group 0.5, and losers go in the group 0). At each round each plays someone with the same score as his/her score. Since the number of perfect scores is cut in half each round, it does not take long until there is only one player remaining with a perfect score.
- Players with similar results play each other as much as possible.
- All players play the same number of matches, regardless of his/her scores.
- Matching of the same rank is realizable.
- Upsets are possible.

Random Swiss Pairing

In Swiss System or its modified version such as Accelerated Swiss, the reverse phenomenon of the merits and qualified counts often occurs. It implies that weaker players may sometimes be finally ranked higher than stronger ones in the sense of Round Robin. In order to improve the drawback, Random Swiss Pairing [3] utilizes Random pairing for some of the rounds and followed by Swiss System. Note that Random Swiss pairing is equivalent to Swiss System when no Random pairing is performed.

2.2. Rating assignment

The winning ratio of higher-ranking player is defined as about 76% with the 200-point difference under the Elo-rating system. To simplify calculation we instead use 75% with the 200-point difference. The ratio W is given by the following formula [3][4], where r represents the rating difference.

$$W = 1 - 1\ /(3^{r/200} + 1) \qquad (1)$$

Below we show the method to determine the win or loss in the experiments.

(1) Obtain the rating difference between two players: White and Black.
(2) In quest of the winning percentage of White a decimal point to the 4th place (rounds off by the 5th place), it doubles 10^4. At this time the value n calculated is in the range of [0, 10000].
(3) The case n = 10000 corresponds to the winning ratio 100% for White, while n = 0 for Black.
(4) Generate a random number r that is in the range of [0, 9999].
(5) If $r < n$, White wins. Otherwise Black wins.

In order to generate the rating, the range which rating can be taken is determined. The rating for each is assigned within the limits. The experiments are conducted with three kinds of rating ranges: 250, 500 and 1000.

The distribution of the rating of players is assumed to be a normal distribution. A distribution is given using the random number that follows a normal distribution to the rating. A normal distribution used the sum of 12 random numbers between 0 and 1. The similar normal random number that practices scaling of this according to the rating range is used.

2.3. Ranking decision

When using Swiss System for a tournament, to eliminate players with the same scores (winning points) the tiebreak system is used [5][6]. For the players who have the same score the ranking is performed in the following order:

(1) A player with the largest number of winning points.
(2) Solkoff (the sum of all the opponents' winning points).
(3) SB (the sum of the winning points of the opponents the player won).
(4) Medium (the sum of the winning points of all the opponents, except the highest and the lowest two opponents).

In this paper we verify which pairing system is superior, as well as their significance based on the measures: winning points, Solkoff, SB and reverse SB (the sum of the number of losses of the opponents the player lost).

2.4. Experimental design

For the experiments we assume that the number of entrants n=20 and rounds t=9. The ranking of all entrants will be assigned. It enables to compare the ranking difference between the rankings by Round Robin and Random Swiss.

It is assumed that Round Robin gives a better ranking than any other pairing systems with limited number of rounds. We also assume that a good pairing system of limited number of rounds becomes closer to the outcome of Round Robin. In the experiments we have a sample of the ranking by Round Robin. Random Swiss takes the rounds of Random Pairing varying from 0 (Swiss System) to 9 (Random Pairing). For one sample ranking by Round Robin, 100 trials are performed by each pairing system concerned and the average is computed. In order to compare the data, "average" and "variance" are calculated to compare each time when randomized [7][8]. Since values with '+' and '-' are intermingled when the ranking difference is compared with Round Robin, "average of difference" and "variance" are calculated by changing them into the absolute values [9]. The formulas are given in (2) - (5).

- Average of difference = Σ (ranking of Round Robin - sampling data) / 100 (2)
- Variance of difference = 1 / (100-1) Σ (sampling data – average of difference)^2 (3)
- | Average of difference | = Σ | (ranking of Round Robin − sampling data) | / 100 (4)
- | Variance of difference | = 1 / (100-1) Σ(| sampling data | − | Average of difference |)^2 (5)

3. Results and Discussions

In the experiments Swiss and Random Swiss are mainly compared. We verify which pairing system (i.e., how many rounds are appropriate for Random pairing in Random Swiss) comes closer to Round Robin, as well as what significance each factor of ranking calculation has.

3.1. Ranking measures

Winning points

The winning points are the number of wins of a player. It does not concern about the strength of the opponents played. The ranking only with the winning points without taking the opponents' strength into consideration may be far from the true ranking of players in the sense of Round Robin.

By comparing the rankings calculated by the number of wins only with that of Round Robin, we observe which system determines effectively by the number of wins only. The procedure is as follows.

(1) The winning points are transformed into ranking.
(2) Tiebreak is not applied.
(3) Calculate the difference between the rankings by Random Swiss and Round Robin.
(4) Calculate "average of the ranking difference" and "variance".
(5) Calculate "average of ranking difference" and "variance" of absolute value.

We show the experimental results for three different rating ranges in Fig.1. Consequently, it turns out that the average and variance of the ranking difference in Round Robin become low as the number of rounds for Random pairing increases, regardless of the rating range. The reason for the possibility that a strong player plays against a weak player becomes high is thought to be from the diminishing opportunities for players of similar abilities to play against each other, resulting in clear outcomes. This is because a win will count as a win, no matter how strong the opponent is. This result appears to reflect the participants' real abilities. However, it is not so reliable since it does not indicate the strengths of the opponents.

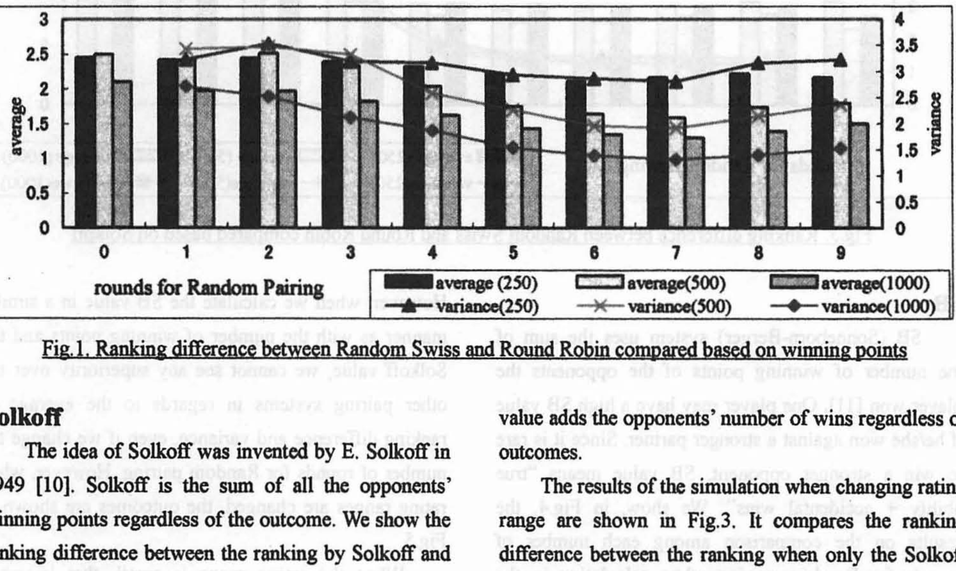

Fig.1. Ranking difference between Random Swiss and Round Robin compared based on winning points

Solkoff

The idea of Solkoff was invented by E. Solkoff in 1949 [10]. Solkoff is the sum of all the opponents' winning points regardless of the outcome. We show the ranking difference between the ranking by Solkoff and Round Robin in Fig.2.

The figure shows the comparison by each number of rounds for Random pairing when setting rating range up to 1000. Horizontal axis is the ranking at the time of Round Robin (finishing sort). A left vertical axis is the average of the ranking difference between Random Swiss and Round Robin. A right vertical axis is a` variance at that time. Fig.2 indicates that as the ranking gets farther away from the middle in the perfect random system, the "average of the difference" with the ranking by Round Robin is larger. This is because the Solkoff value adds the opponents' number of wins regardless of outcomes.

The results of the simulation when changing rating range are shown in Fig.3. It compares the ranking difference between the ranking when only the Solkoff value is used and the ranking by Round Robin. It turns out that the ranking difference becomes large as the number of rounds for Random pairing increases. This is because as the number of rounds for Random pairing increases, many matches are assigned at random, regardless of their abilities, creating bias pairing. Furthermore, since the Solkoff value does not reflect players' abilities, it creates a wider gap in the number of wins between high/low ranked players. This indicates that it is important for the Solkoff system to set up matches with opponents' abilities in mind.

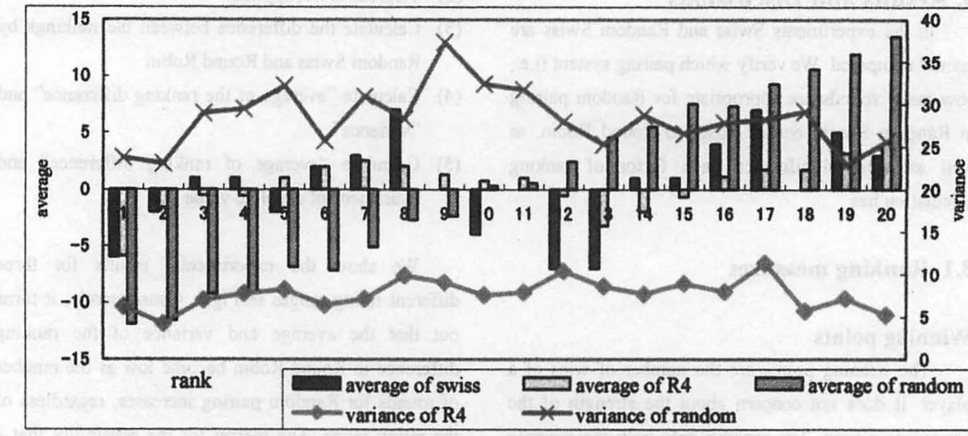

Fig.2. Random Swiss, Round Robin and Solkoff ranking compared

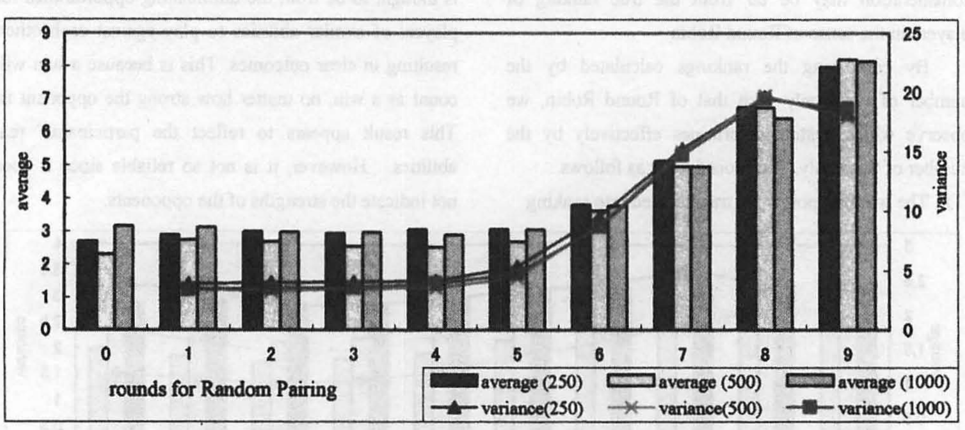

Fig.3. Ranking difference between Random Swiss and Round Robin compared based on Solkoff

SB

SB (Sonneborn-Berger) system uses the sum of the number of winning points of the opponents the player won [11]. One player may have a high SB value if he/she won against a stronger partner. Since it is rare to win a stronger opponent, SB value means "true ability + accidental wins". We show, in Fig.4, the results on the comparison among each number of rounds for Random pairing when calculating in the absolute value at case rating range 1000.

The results of Fig.4 show that a variance value also becomes high when the average of ranking difference between Random Swiss and Round Robin is high. An accidental wins of SB value influences this, therefore, the average of ranking difference becomes high. Since the act of being accidental is a rare occurrence, the variation from the part average is also considered to be because it also has become high. However, when we calculate the SB value in a similar manner as with the number of winning points and the Solkoff value, we cannot see any superiority over the other pairing systems in regards to the average of ranking difference and variance, even if we change the number of rounds for Random pairing. However, when rating ranges are changed, the outcomes are shown in Fig.5.

When the rating range is small, that is, when abilities are balanced, it becomes easier for upsets to take and as the result. SB value becomes high. However, when there is a clear difference in rating range, indicating a clear difference in players' abilities, then upsets is less likely to occur. Therefore, SB value is thought that it correlates with abilities, such as with rating range, and not correlated with a pairing system like the number of rounds for Random pairing.

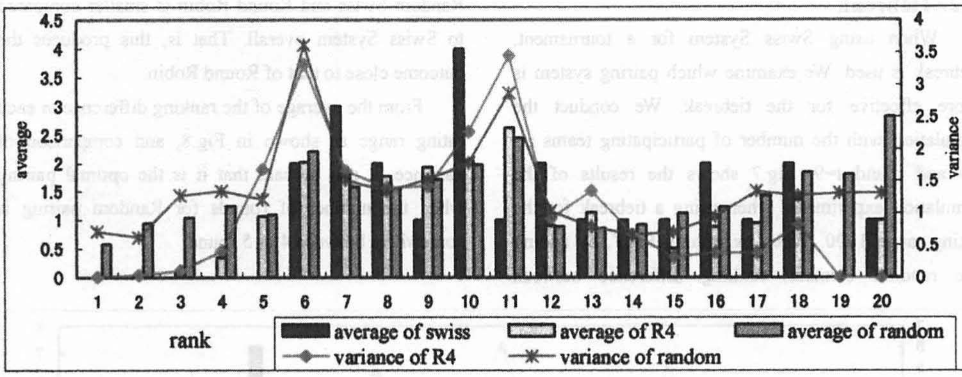

Fig.4. Random Swiss, Round Robin and SB ranking compared

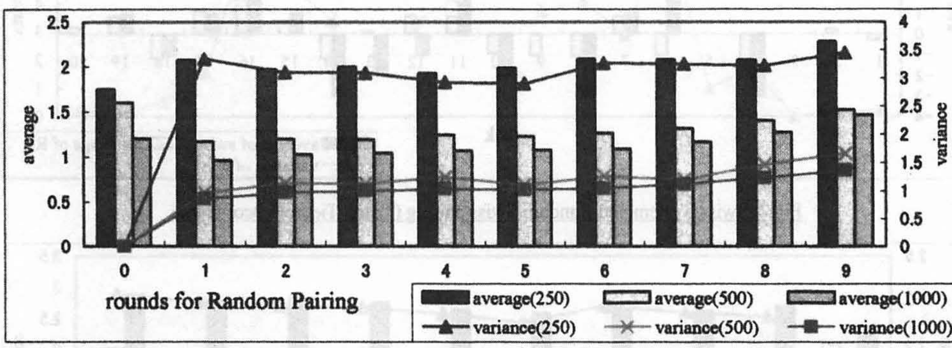

Fig.5. Ranking difference between Random Swiss and Round Robin compared based on SB

SB enhancement

SB concerns "True ability + accidental wins". We can also concern the sum of the number of losses for those who you lost. Namely we also consider the accidental losses by weaker opponents. This is "True ability + accidental losses" (called the reverse SB hereafter). The results of the simulation with rating range of 250 are shown in Fig.6.

Fig. 6 shows the difference in results with reverse SB with ranking determined by fewer numbers of losses, compared to that by Round Robin. When analyzing Fig. 6, we can see that when the average of ranking difference between Random Swiss and Round Robin is high, it turns out like the original SB value that a variances value also becomes high. One can consider that it is not necessary to see this type of data since it is the same as SB value. However, when a game is played with the limited number of rounds, the negative portion, indicating whether you have lost by chance, as well as the how weak your opponent was should also be examined.

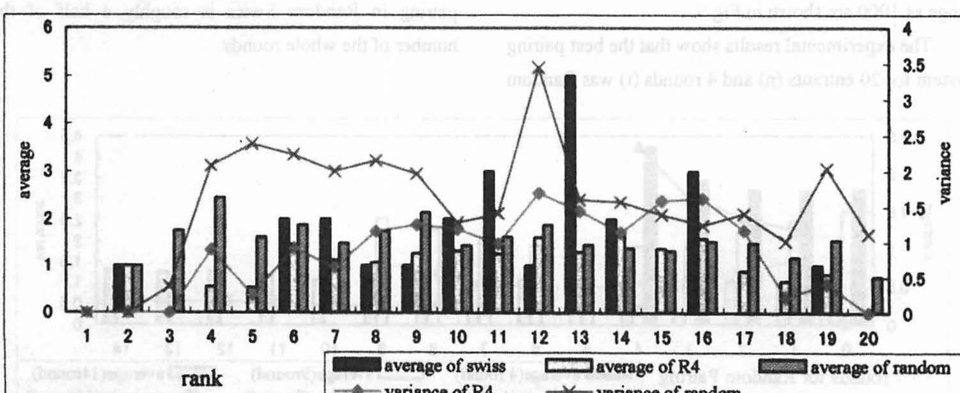

Fig.6. Random Swiss, Round Robin and reverse SB ranking compared

3.2. Tiebreak

When using Swiss System for a tournament, tiebreak is used. We examine which pairing system is more effective for the tiebreak. We conduct the simulation with the number of participating teams n= 20 and round t=9. Fig.7 shows the results of the simulation experiments when using a tiebreak for the rating range 1000. When examining Fig.7, by adding the random element, ranking difference between Random Swiss and Round Robin is smaller compared to Swiss System overall. That is, this produces the outcome close to that of Round Robin.

From the average of the ranking difference in each rating range as shown in Fig.8, and comparison of variance, it can be said that it is the optimal pairing when the number of rounds for Random pairing is somewhere between 4 or 5 round.

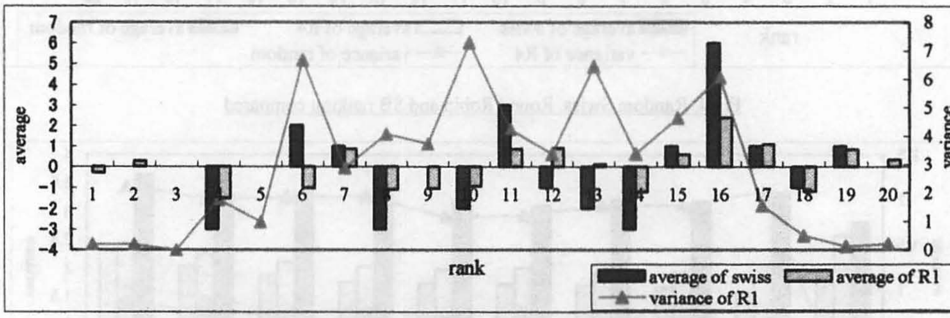

Fig.7. Swiss System and Random Swiss pairing (using Tiebreak)compared

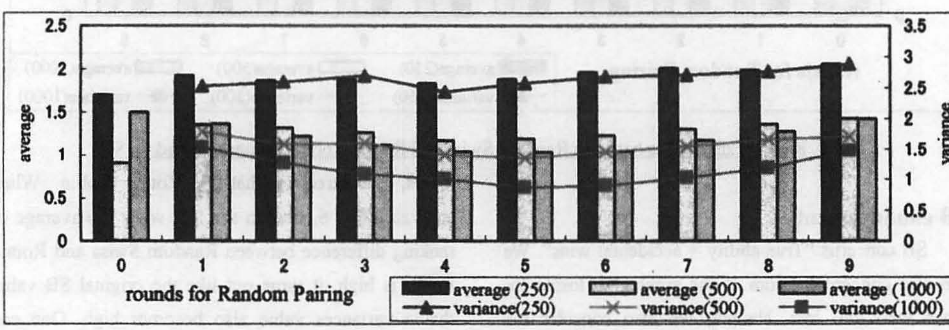

Fig.8. Random Swiss and Round Robin compared for various rating range

3.3. Rounds for Random pairing

We conduct the simulation with the number of entrants (n) fixed to 20 teams. The rounds per team (t) are set to 4, 9, and 14. The results when setting rating range as 1000 are shown in Fig.9.

The experimental results show that the best pairing system for 20 entrants (n) and 4 rounds (t) was Random Swiss with Random pairing twice. Random Swiss with 4 or 5 rounds for Random pairing in case t=9 and 7 rounds for Random pairing in case t=14. These results show that the optimal number of rounds for Random pairing in Random Swiss is roughly a half of the number of the whole rounds.

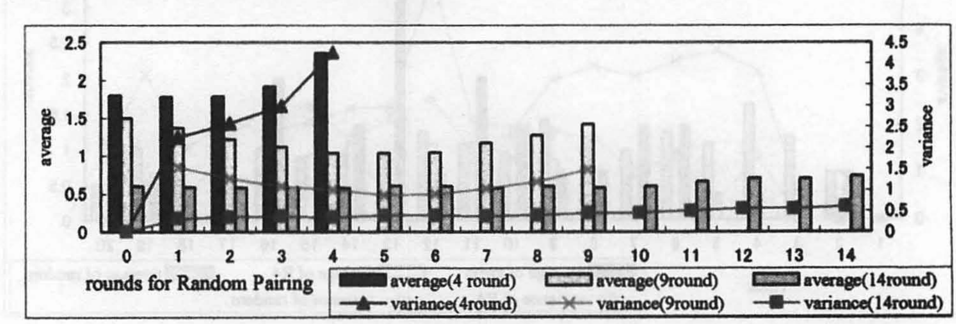

Fig.9. Ranking difference between Random Swiss and Round Robin with various rounds (t) for Random Pairing

3.4. Number of tournament entrants

What will happens when increasing the number of entrants for the fixed number of rounds? For a tournament of 9 rounds we set four different entrants n=24, 40, 60, and 80. We show, in Fig.10, the results in the case where the rating range is 500. Fig.10 indicates that it does not make any difference even though Random-pairing rounds are included. There are only few rounds played as the number of entrants increase. This is because the ranking takes place after early rounds and ignoring the results thereafter. The number of the minimum required rounds for Knockout tournament is given by Equation (6).

$$X = [\log_2(n)] \quad (X: \text{minimum required rounds, n : entrants}) \quad (6)$$

We show, in Table 1, our calculation for the minimum required number of rounds in a Knockout tournament for various numbers of entrants.

Entrants (n)	24	40	60	80
Minimum required rounds (X)	5	6	6	7

Table.1. The minimum required rounds in Knockout tournament and the numbers of entrants.

The number of rounds for Random pairing in Random Swiss is given as the difference between the number of all rounds played and the minimum required rounds for Knockout. It may suggest a sufficient number of all rounds for a tournament when applying Random Swiss since Swiss System possesses the property of Knockout tournament. We simply recognize that the number of the rounds is twice of the minimum required rounds for Knockout.

$$Y = X \times 2 \quad Y: \text{the minimum required rounds played in a tournament,}$$
$$X: \text{the minimum required rounds for Knockout} \quad (7)$$

Experiments are conducted to verify the relation of Formula (7) based on the minimum required rounds shown in Table 1. We show the results in Fig.11. In the experiment the number of all rounds for a tournament is set with the twice number of rounds for the minimum required rounds for Swiss System. The results show that the accuracy improves Swiss System. Naturally, if the number of rounds increases, the accuracy improves as well. However, it may be ideal if we do not conduct all the matches by Swiss System. Instead it might be better to pair as many matches with the random system as the Swiss System.

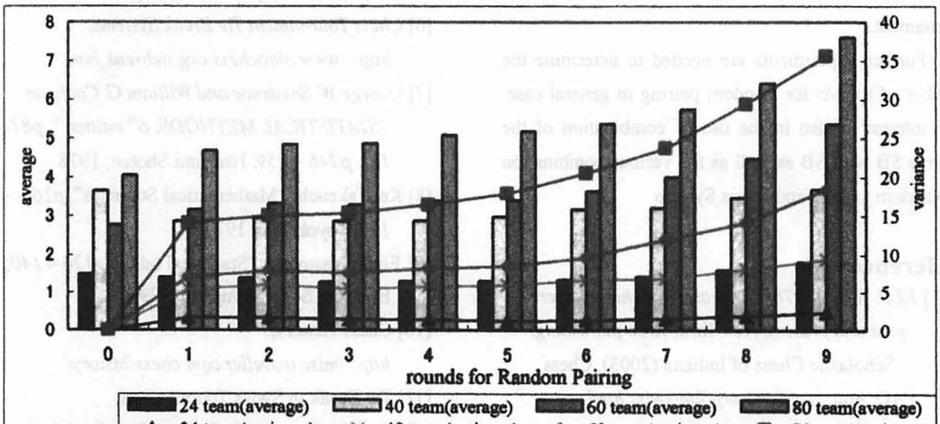

Fig.10. Ranking difference between Random Swiss and Round Robin with various numbers of entrants

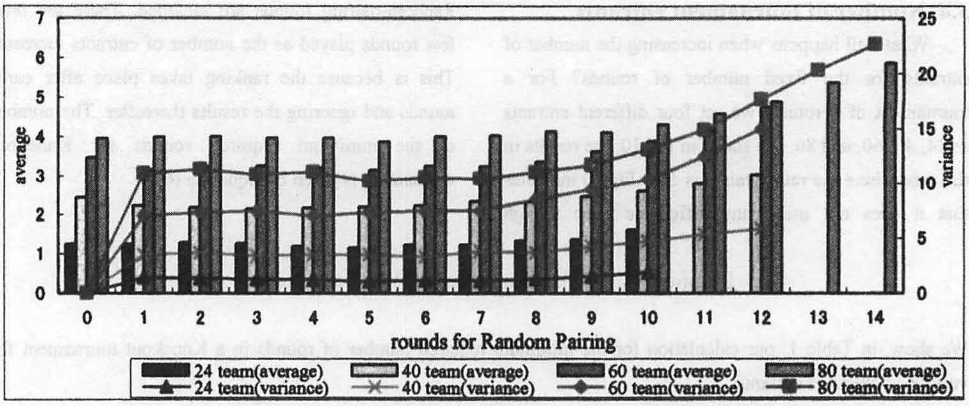

Fig.11. Ranking difference and the number of Random pairing rounds: the minimum required rounds performed for various number of players.

4. Conclusion

In this paper we studied on pairing systems that produce the rankings for tournament entrants that would reflect the real strength of each player even when the number of rounds is restricted. In the experiments Random Swiss pairing, with various number of rounds for Random pairing, was compared with Round Robin in the agreement of the ranking based on the measures such as winning points, Solkoff, SB and reverse SB. The significance of each measure for the ranking was also considered.

It is found that Random Swiss pairing outperforms in any case Swiss System. The optimal number of rounds for Random pairing in Random Swiss is roughly a half of the number of the whole rounds. We suspect that the number of the rounds necessary for an effective Random Swiss tournament is twice of the minimum required rounds to select the winner in a Knockout tournament.

Further experiments are needed to determine the number of rounds for Random pairing in general case. Our interest is also in the use of combination of the reverse SB with SB as well as the various combination of Random pairing and Swiss System.

References

[1] *ELO, A. E. (1978). The rating of chess players past and present, New York: Arco publishing.* Scholastic Chess of Indiana (2003). Chess FAQ. *http://scichess.org/faq/swiss.html*

[2] Takenobu TAKIZAWA and Yoshikazu KAKINOKI "A Pairing System and Its Effectiveness in the World Computer Shogi Championship" , The 7th Game Programming Workshop in Japan 2002 (*IPSJ Symposium Series Vol.2002, NO.17*) *p93-100*

[3] Tsuyoshi Hashimoto, Jun Nagashima, Hiroyuki Iida, "A Proposal of tournament system verification by the simulation – a case study using the World Computer Shogi Championship -", The 7th Game Programming Workshop in Japan 2002 (*IPSJ Symposium Series Vol.2002, NO.17*) *p101-107*

[4] About rating Shogi, Shogi Club 24 homepage, *http://www.shogidojo.com/dojo/rating.html*

[5] Computer Shogi Championship Enforcement point, Computer Shogi Association homepage, *http://www.computer-shogi.org/*

[6] *Chess Tournament Tie Break Systems, http://www.ohiochess.org/tiebreak.htm*

[7] *George W. Snedecor and William G. Cochran: "STATISTICAL METHODS, 6th edition", p87 – 112, p246 – 259,* Iwanami Shoten, 1978

[8] Kei Takeuchi "Mathematical Statistics", *p165 – 170,* Toyokeizai, 1977

[9] Eiichi Kasuya: "Statistical talk", *p126 – 140,* Bun-ichi Sogo Shyuppan, 1999

[10] Chess History, *http://misc.traveller.com/chess/history/*

[11] Tie-Break in Swiss Tournaments, *http://www.swissperfect.com/handbook/tiebreak.htm*