

トーナメントスコアに基づく順位決定の新たな指標

河合 孝尚¹, 川島伸太郎², 飯田弘之^{3,4}

1. 静岡大学工学部理工学研究科 2. 静岡大学情報学部
3. 静岡大学大学院情報学研究科 4. 科学技術振興事業団さきがけ研究 2 1

概要

競技参加者の順位付け（例えば、メダル受賞者の選抜およびその基準など）に関する問題は、競技者の実力を正当かつ公平に評価するという観点から実に重要な課題である。本稿では、従来の順位計算で用いられてきたソルコフ値およびSB値から得られる新たな指標を提案する。新しい指標を用いた順位決定方式と従来の方式（単純に勝数の多い順に順位を付ける）を比較・評価するために、シミュレーション実験を実施した。実験の結果、本稿で提案する指標を用いた順位決定方式のほうが従来方式よりも、総当り方式の結果により近くなるという観点で優れていることを確認した。

A new ranking measurement based on tournament records

T. Kawai¹, S. Kawashima² and H. Iida^{3,4}

1. Graduate School of Science and Engineering, Shizuoka University
2. Faculty of Computer Science, Shizuoka University
3. Department of Computer Science, Shizuoka University
4. PRESTO, Japan Science and Technology Agency

Abstract

It is an important issue to evaluate duly and impartially the tournament performance of the participants in any competitions, especially in the case where medal winners have to be distinguished. This paper proposes a new measurement obtained based on Solkoff and SB points of a tournament. Simulation experiments were performed to compare the ranking system using the proposed measurement with the traditional ranking system that is based on the winning point, Solkoff and SB points. The experimental results show the effectiveness of the proposed idea in the sense that the proposed ranking system gives a ranking closer to the round robin system.

1. はじめに

オリンピックのような肉体的競技会からチェスのような知的競技会まで、様々な種類かつ規模の競技会が世界各地で多数開催されている。その際、プレイヤー自身の実力を判定するために様々な規模の競技大会が世界各地で開催され、その参加者の中で実力の優劣を決定するために順位計算が行なわれている。しかし、どの順位計算方法がより実力を正確に反映させているのか定かではない。しかも開催期間や組み合わせ方式などの制約の範囲内で大会が行なわれるのでプレイヤーの真の実力を測ることは非常に困難な問題である。そこで本稿では、今までの勝数の多い順に順位を付けるのではなく、順位計算で用いられているソルコフ値、SB値を利用して新しい指標を考案し、この指標を用いた組み合わせ方式と、勝数の多い順に順位を付けた組み合わせ方式を比較することで、どちらが自分以外の参加者全員と対戦する総当り方式に近いのかシミュレーション実験を行なった。

2. シミュレーション

参加チーム数と試合数が与えられたとき、大会の方式はより実力が反映される方式、つまり強いものがより多く勝ち上がる方式が妥当であると考えられる。そのためには各チームにレーティングを与え、その各チームのレーティングに応じて試合の勝敗を確率的に決めればシミュレーションを行なうことが可能となる。これにより各組み合わせ方式について何パターンかシミュレーションを行なうことができ、順位の妥当性や対戦の偏りについて検証することが可能となるのである。それにより大会の組み合わせ方式の妥当性を考える上での重要な材料になるものだと考える。

2.1. 組み合わせ方式

今回の実験では、以下の組み合わせ方式を使用し、シミュレーション実験を行なった。

2.1.1. 総当り方式

総当り方式では、全競技者が他の全競技者と対戦する。現在、参加競技者全員に順位をつけなければならない競技会で、最も公平な競技方式は総当り方式だと考えられている。しかし、競技者数が多いと膨大な時間を要するので、数多くの競技者の順序付けをある期間内に決定しなければならない場合など、現実的には実行不可能であることも多い。そこで、実際にはノックアウト方式やスイス方式等の限定ラウンド方式がよく用いられる。また、これらの限定ラウンド方式の公平性を評価する際、総当り方式の結果にどれほど近いかで、公平性を比較・評価することができる。

2.1.2. スイス方式

スイス方式は、囲碁、将棋、連珠、チェスの大会などで利用されている変則リーグ戦方式である[1]。ノックアウト方式では、フロックによる勝敗が大きな要素を占めてしまう。敗者復活方式やシード方式を導入することもあるが、それでも実力が正確には反映され難い。スイス方式はその欠陥を補うために考え出された方式であり、従来、限定ラウンドの試合回数を実施せざるを得ない場合、スイス方式が伝統的に信頼される方式として用いられてきた。

この方式では、競技会進行プロセスで同じもしくは近いスコアの競技者同士を対戦させる。ただし、同一競技者との対戦は一度限りとする。たくさんの参加競技者がいるとき、少ない試合ラウンド数で、かつ、各競技者とも同じラウンド数の試合を行なう。主な特徴を以下に列挙する。

- (a) その時点で同じかできるだけ近いスコアの競技者と対戦する
- (b) スコアに関わらず全競技者が同じラウンド数の試合に参加する
- (c) 競技会途中での棄権は認めない
- (d) 順位逆転の余地がある
- (e) 対戦の組み合わせが非常に煩雑である

また、近年、私たちはスイス方式を改善する方式(ランダムスイス混合組み合わせ方式)を考案した[2]。この方式では、複数の参加チームを2チームずつ互いに対戦させる対戦型の競技会等の場合において、対戦の組み合わせによっては弱いチームが強いチームよりも上の順位になってしまうという従来の問題を解決し、公平性および順位の妥当性を確保しているものである。

2.2. 順位計算方式

今回の実験では、以下の順位計算方式を使用しシミュレーション実験を行なった。

2.2.1. FIDE方式

従来よく使われる順位計算の方法は、勝数の多い競技者から順に順位付けすることである。ここで、同率順位をなくすために、FIDE(世界チェス連盟)が採用している順位決定法が従来よく知られている[3]。この方法では、競技会の対

戦組み合わせ方式としてスイス方式が用いられている場合、同率順位をなくすタイブレークのための指標として、勝敗スコアから以下の4つの指標を用いる。

- (a) 勝数
- (b) ソルコフ方式による計算値(対戦相手全員の勝数の合計)
- (c) SB方式による計算値(対戦して負かした相手全員の勝数の合計)
- (d) ミディアム方式による計算値(負かした相手の勝数が最高と最低の2人を除いた相手の勝数の合計)

そして、(a)の勝数が同じ(同順位の)競技者が複数いる場合、(b)ソルコフ値の高い競技者、(c)SB値の高い競技者、(d)ミディアム方式による値が高い競技者、という優先順序で順位付けを行なう。

2.2.2. ソルコフ方式

ソルコフ値とは、対戦相手全員の勝数の合計である。すなわちソルコフ値は「対戦相手の見かけ上の強さ」を表していると考えることができ、強い(勝数の多い)対戦相手と組み合った競技者のソルコフ値は高く、弱い(勝数の少ない)対戦相手と組み合った競技者のソルコフ値は低くなる。

2.2.3. SB方式

SB値とは、対戦して負かした相手全員の勝数の合計である。したがって、自分より強い対戦相手に勝った競技者のSB値は高くなり、自分より弱い対戦相手に負けた競技者のSB値は低くなる。すなわちSB値は、「競技会での番狂わせ」を表すものと考えられる。

2.3. 新指標について

従来のFIDE方式では、勝数が同じ(同順位の)競技者が複数いる場合に、ソルコフ値の高い競技者、SB値の高い競技者、という優先順序で順位付けを行なう。これに対し、本研究では次の指標Xを用いた順位付け方法を提案する。

$$X = SB / SOL \quad \dots (1)$$

なお、SOLはソルコフ値(対戦相手全員の勝数の合計)、SBはSB値(対戦して負かした相手全員の勝数の合計)である。そして、指標Xが高い順に上位から順位付けする。なお、競技者のXが最も高くなるのは、その競技者が対戦に全勝した場合であり、その値は1となる。この式(1)による指標は、ある競技者が、どのような相手と試合をして、どのように勝数を獲得したかを示している。従って、この方法により順位付けを行なえば、対戦の偏りに影響されることなく、より競技者の真の実力を測ることができるといえる。

また、次式(2)に示すように、勝数(win)に上式(1)の値を指数として掛けた指標を用いても、同様の効果が得られる。

$$X = win^{(SB/SOL)} \quad \dots (2)$$

式(2)においても、指標Xが高い順に、上位から順位付けする。また式(2)では指数部分の(SB/SOL)が0であっても、X=1になるので、下位の順位のプレイヤーに差をつけるものとも考えられる。

2.4. 勝敗判定について

勝敗判定についてはEloレーティングを使用する[4]。Eloレーティングでは200点差で上位者の勝率は約76%と定義されているが、計算を簡単にするため、ここでは200点差で75%とする。上位者の勝率は以下の式で計算する[4][5]。

$$\text{上位者の勝率} = 1 - 1 / (3^{\text{レーティング差} / 200} + 1)$$

勝敗の判定の仕方は以下のようにする。

- (1) 先手と後手のレーティング差を求める
- (2) 先手の勝率を小数点以下第4位まで求め(第5位で四捨五入) 10^4 倍する。このとき、求められる値(nとする)は0~10000である
- (3) nが10000の時(=先手が勝率100%)は先手の勝ち,0の場合は後手の勝ちとする
- (4) 乱数rを0~9999の範囲で発生させる
- (5) $r < n$ なら先手勝ち, $r > n$ または $r = n$ なら後手勝ちとする

3. 実験1

本稿では、順位決定システムを用いて2つのシミュレーション実験を行ない、本研究の評価を行った。

3.1. タイブ레이크なし(勝ち,負け)の順位計算

3.1.1. シミュレーションの方法と手順

第1のシミュレーションは、まず、参加チーム数を20として、総当り方式とスイス方式の2つの対戦組み合わせ方式を用いて行なう。スイス方式については、試合ラウンド数を9回戦とした。なお、ここでは限定ラウンド方式の一例としてスイス方式を用いたが、他の方式を用いてもよい。また、シミュレーションの前提として、各チーム(Team00~Team19)に強さ(レーティング)を与える。各チームのレーティングはTeam00を最大値にし、そこから降順で与えている。そして総当り方式とスイス方式の2つの対戦結果を入力データとして、以下の方法で実験を行なった。

- (1) 総当り方式で全チームを対戦させる
- (2) 試合結果は勝ち(2点),負け(0点)とする
- (3) 任意の指標に基づいた順位付けを行う(FIDE方式,新指標方式)
- (4) スイス方式での対戦を行う
- (5) 任意の指標に基づいた順位付けを行う
- (6) 各チームの順位について、総当り方式とそれぞれの順位計算で求めた順位との差を求める

3.1.2. シミュレーションの結果と考察

シミュレーション結果を図1に示す。

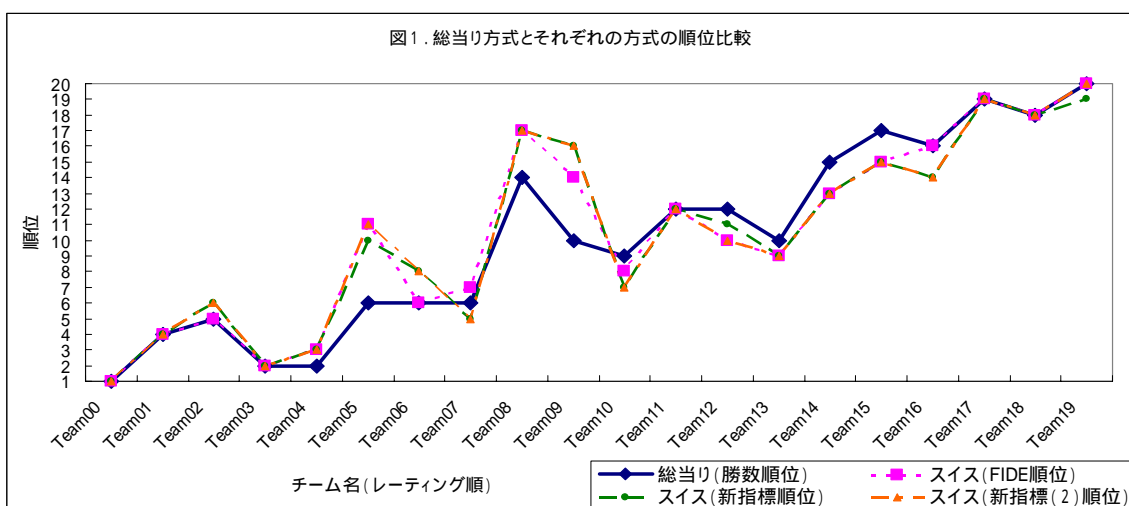


図1. 総当り方式とそれぞれの順位計算方式での順位比較

この結果、ほとんどFIDE方式と同等の順位計算結果が得られることがわかった。もともとFIDE方式という順位計算方式は多くの大会などで使用されており、すでに高いレベルで完成された順位計算方式といえるので、それと同等の結

果が得られたということはそれなりに評価できるものだと考えられる。しかし、ここでの実験では新指標(1)($X = SB/SOL$)と新指標(2)($X = win^{(SB/SOL)}$)での優位さは見られなかった。

3.2. タイブ레이크を含めた(勝ち, 負け, 引き分け)順位計算

3.1の第1のシミュレーション実験においては、引き分けが存在しないものとしてシミュレーションを行なった。しかしながら、実際には、チェスなど引き分けの多い競技は多数存在する[6][7]。そこで、引き分けを考慮した第2のシミュレーションも行なって、同様に本研究の効果を示す。

3.2.1. タイブ레이크について

はじめに、各チームに与えたレーティングを用いて総当たり方式による対戦を行ない、勝敗引き分けの結果を出す。勝敗の決定には、3.1の第1のシミュレーションと同じく Elo レーティングを用いる。ここで、Elo レーティングだけでは引き分けを求められないので、以下の手順で勝敗引き分けを決定する[4]。

(1) 先手の勝率を計算

(2) 乱数を発生

(3) 引き分け率を計算して乱数と比較

ここで、引き分け率を求めるには従来技術である

$$P_{\frac{1}{2}} = \sqrt{\frac{P_{ij}(1-P_{ij})}{2}}$$

の式を用いる。ここで、 P_{ij} は、プレイヤー i がプレイヤー j に勝つ確率である。

(4) 上記(3)の結果が引き分けの場合はその対戦結果を引き分けとし、引き分けではない場合は、再び乱数を発生させ Elo レーティングを用いて勝敗の決定を行なう。

3.2.2. シミュレーションの方法と手順

第2のシミュレーションとして、まず、参加チーム数を24として、総当たり方式とスイス方式の2つの対戦組み合わせ方式を用いて行なった。スイス方式については、試合ラウンド数を9回戦とした。シミュレーションの前提として、各チーム(Team00~Team23)に強さ(レーティング)を与える。各チームのレーティングはTeam00を最大値にし、そこから降順で与えている。そして総当たり方式とスイス方式の2つの対戦結果を入力データとして、以下の方法で実験を行なった。

(1) 総当たり方式で全チームを対戦させる

(2) 試合結果は勝ち(2点)、引き分け(1点)、負け(0点)とする

(3) 任意の指標に基づいた順位付けを行う(FIDE方式、新指標方式)

(4) スイス方式での対戦を行う

(5) 任意の指標に基づいた順位付けを行う

(6) 各チームの順位について、総当たり方式とそれぞれの順位計算で求めた順位との差を求める

3.2.3. シミュレーションの結果と考察

シミュレーション実験によって求めた総当たり方式とスイス方式の各順位計算での計算結果をそれぞれ比較したものを図2と図3に示す。

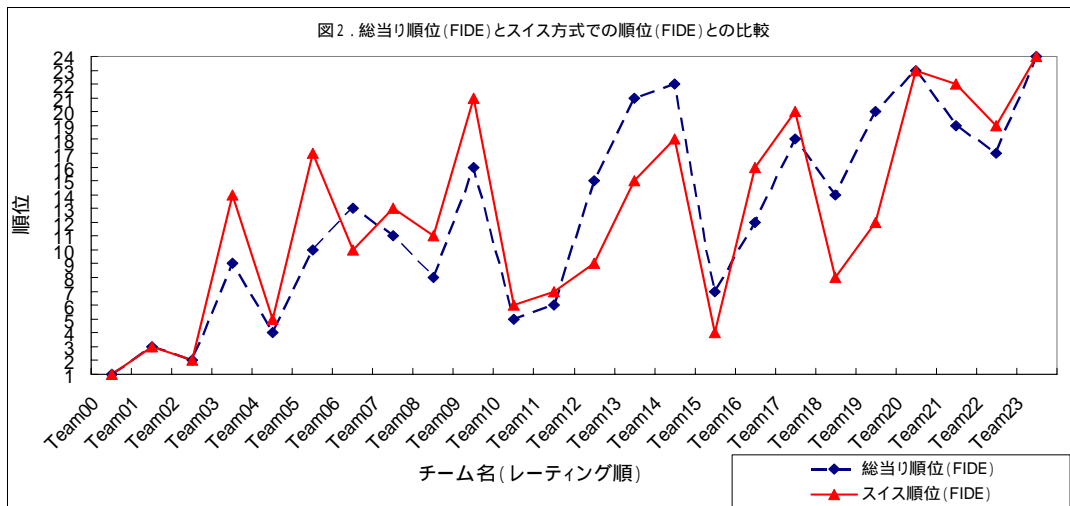


図2 . 総当り方式とスイス方式のF I D E方式での順位比較

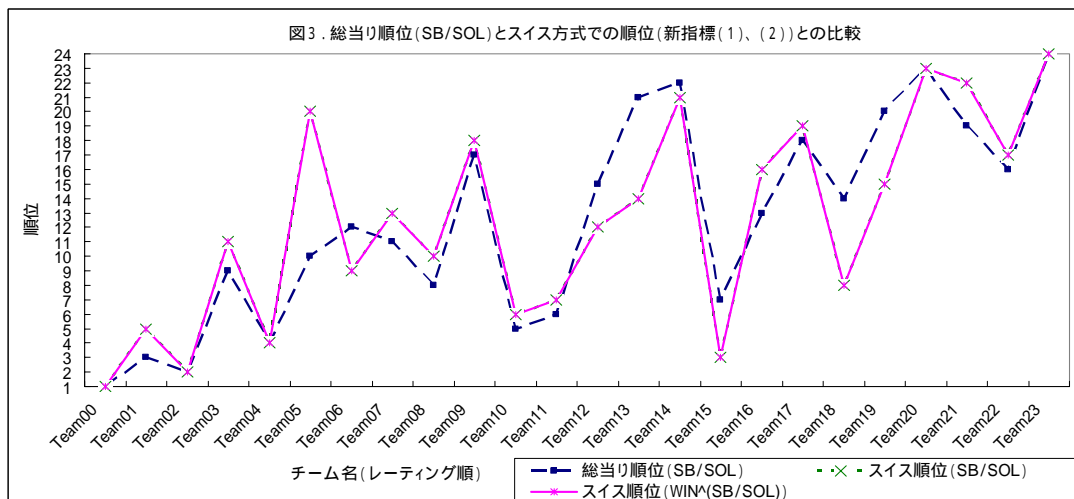


図3 . 総当り方式とスイス方式の新指標方式(1), (2)での(2)順位比較

ここで、図2と図3のそれぞれの総当り方式との順位差を求めたものを図4に示す。

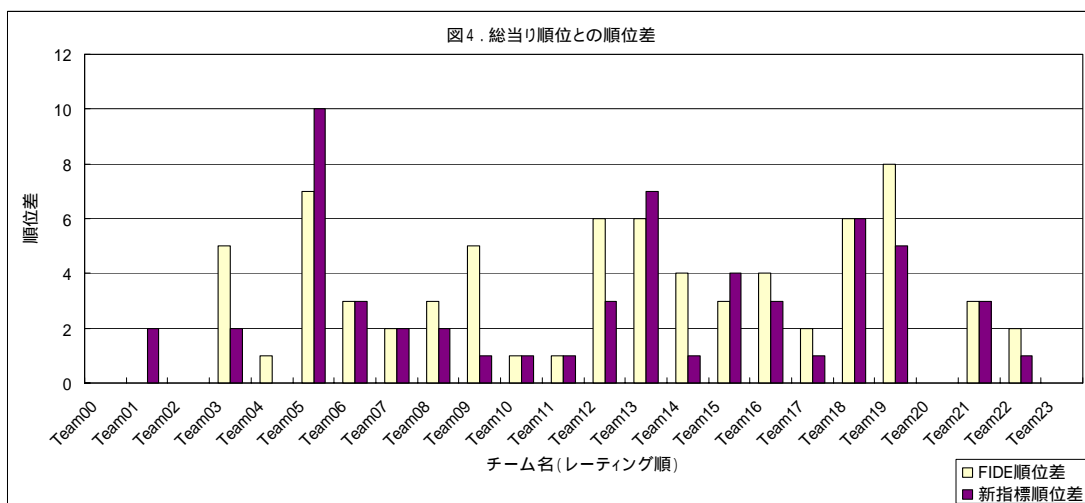


図4 . チーム別 (レーティング別) の各順位計算での順位と総当り方式の順位との順位差

図4より、F I D E方式での総当り方式との順位差の合計は72となり、一方、新指標方式（図3より新指標(1)と(2)は同値だったので新指標方式と省略する）では順位差の合計は58となった。つまり、新指標方式のほうがF I D E方式よりも、より総当り方式の結果に近いといえる。総当り方式の順位との一致回数、つまり順位差が0の箇所はF I D E方式、新指標方式ともに5チームであった。

4. 実験2

ここまでの実験では1つの総当り表から1つのスイス方式での組み合わせについてだけ論じてきた。しかし、スイス方式での組み合わせは限定試合回数での対戦結果によって様々に見られる。そこで次に、1つの総当り表からいくつかのスイス方式での組み合わせを行い、それぞれのチームに与えたレーティングの幅を変化させて実験を行なった。

4.1. シミュレーションの方法と手順

参加チーム数を20として、総当り方式1回に対し、スイス方式の組み合わせを100回行なった。スイス方式については、試合ラウンド数を9回戦とした。シミュレーションの前提として、各チーム(Team00~Team19)に強さ(レーティング)を与える。各チームのレーティングはTeam00を最大値にし、そこから等間隔で降順に与えている。そして総当り方式とスイス方式の2つの対戦結果を入力データとして、以下の方法で実験を行なった。

- (1) 総当り方式で全チームを対戦させる
- (2) 試合結果は勝ち(2点), 引き分け(1点), 負け(0点)とする
- (3) 任意の指標に基づいた順位付けを行う(FIDE方式, 新指標方式)
- (4) スイス方式での対戦を行う
- (5) 任意の指標に基づいた順位付けを行う
- (6) 各チームの順位について、総当り方式とそれぞれの順位計算で求めた順位との差を求める
- (7) (4)~(6)を100回繰り返す。
- (8) 100回求めた順位差の平均と分散を求め、それに順位を付け、総当り方式の順位と比較する
- (9) 総当り方式との順位差を求め、それを合計する
- (10) 順位差が0のものをカウントする(一致回数)

4.2. シミュレーションの結果と考察

シミュレーションの結果を図5に示す。

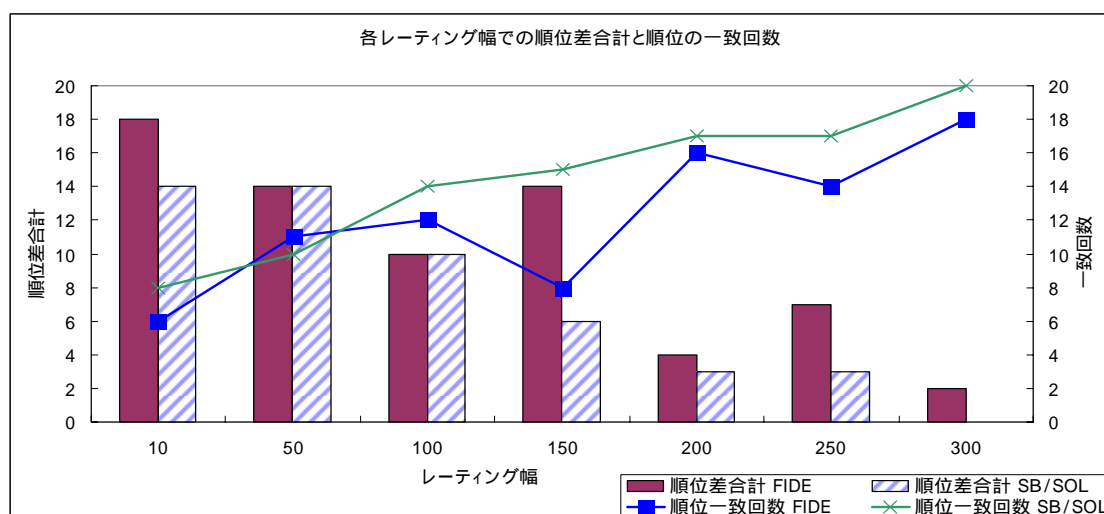


図5. 各レーティング幅での順位差合計と順位的一致回数

図5の結果から、F I D E方式のほうが順位にばらつきがあることがわかった。通常、レーティング幅が小さいときはどちらのチームが勝つかわからないので順位差の分散は大きくなり、そのため合計値も大きくなる。しかしレーティング幅が大きくなると勝負の勝ち、負けがはっきり出るので、順位差の合計は小さくなる。図5を見ると、新指標方式のほうはレーティング幅が大きくなるにつれて順位差の合計も小さくなっているのに対し、F I D E方式ではレーティング幅によって順位差の合計にばらつきがみられた。一方、順位差が0（総当り方式の順位と一致）のチーム数は、新指標方式のほうはレーティング幅があがるにつれ一致回数も増えているのに対し、F I D E方式では、またもやばらついていた。このことから、新指標方式のほうが、どんな対戦回数、対戦組み合わせが行なわれても対応できることがわかった。

5 . 結論

本稿では、今までの勝数の多い順に順位を付けるのではなく、順位計算で用いられているソルコフ値、SB値を利用して新しい指標を考案し、この指標を用いた組み合わせ方式と、勝数の多い順に順位を付けた組み合わせ方式を比較することで、どちらが自分以外の参加者全員と対戦する総当り方式に近いのかシミュレーション実験を行なった。その結果、一般的に用いられている勝数の多い順に順位を決定するよりも、新指標（SB/SOL）を用いた順位計算のほうが、より総当り方式の結果に近づくことができることがわかった。つまり、スイス方式という実用的な対戦方式下において、今までのような勝数に基づいた順位付けと、新指標（SB/SOL）値に基づいた順位付けを比較し、どれだけ総当りの結果に近づくかシミュレーション実験を行なった結果、新指標に基づく順位付けのほうがより総当りの結果に近く、すなわち順位付けの指標として精度が高いと言える。

6 . 参考文献

- [1] 瀧澤 武信, 柿木 義一 “世界コンピュータ将棋選手権における対戦組み合わせシステムの有効性（2）”,
The 8th Game Programming Workshop in Japan 2003, p122 - 125 (2003)
- [2] T. Kawai, H. Iida: “Random Swiss Pairing System”,
The 8th Game Programming Workshop in Japan 2003, p137 - 144 (2003)
- [3] 第12回世界コンピュータ将棋選手権実施要領 <http://www.computer-shogi.org/wcsc12/youryou.html>
- [4] T. Hashimoto, J. Nagashima, H. Iida: “A proposal of tournament system verification by the simulation
- a case study using the World Computer Shogi Championship -”,
The 8th Game Programming workshop in Japan 2002, p101 - 108 (2002)
- [5] レーティング将棋について, 将棋倶楽部24 ホームページ, <http://www.shogidojo.com/dojo/rating.htm>
- [6] ELO, A. E. The rating of chess players past and present, New York: Arco publishing. (1978)
- [7] Scholastic Chess of Indiana (2003). Chess FAQ. <http://scichess.org/faq/swiss.html>