AN ADJUSTIVE RATING SYSTEM FOR TENNIS
AND SQUASH PLAYERS

Stephen R. Clarke¹

Abstract

The advantages of a rating system for tennis and squash players, incorporating elite down to club players, are discussed. Some problems with rating systems currently in use are highlighted. The use of exponential smoothing as a rating method is proposed, and its possible application to tennis and squash demonstrated. Possible solutions to some minor difficulties are suggested.

1. INTRODUCTION

Imagine my difficulties in organising a tennis match while I am at the conference.

Steve: Bill, would you like a game of tennis?
Bill: I might be interested, but how good are you?
Steve: Well I play section A1 mixed on Thursday night Blackburn district competition, but section A4 on Saturday Eastern Suburban competition. Years ago I played No 3 in C special pennant.

Bill: I play veterans pennant section 2 in Victoria, but I have no idea how that compares - we will just have to hope we are the same standard.

As a contrast consider golf:

Bill: My handicap used to be 1, but over the last few years I have dropped to 5. What's yours?
Steve: I don't have an official handicap, but I play to a handicap of about 28. I heard that American say he is off 3, why don't you try him for a game?

Golf has a recognised system of rating all players from different courses, states and countries. The system uses the same scale for the best golfers in the world down to the weekend hacker. A golfer can always strive to improve or hold his handicap. He can compare his current handicap with his best-ever handicap. He can compare his best with another player's best, even if they played in different eras.

On the other hand, tennis players have trouble determining relative standards even within a club. It would be nice to have a system that rated tennis or squash players, from the ordinary club player to the best in the world, in a similar manner to the handicap system for golf.

This article broaches the possibility of such a rating and suggests some possibilities.

¹Mathematics Department, Swinburne University, Hawthorn, Vic.
2. A CURRENT SQUASH RATING SYSTEM

Various rating systems already exist in racquet sports. There is of course a computer tennis ranking system for elite players. Players accumulate points depending on how far they progress in each tournament, but not on whom they defeat and the margin of victory. A win over a strong opponent carries the same reward as a win over a weak opponent. A close loss to a strong opponent earns no more points than a thrashing by a weak opponent. Stefani [1] calls this a cumulative scheme, as distinct from an adjutative scheme whereby players' ratings are adjusted up or down depending on their performance relative to that predicted by their rating. Hoffman [2] uses the terms 'performance' to describe what is measured by a cumulative scheme, and 'ability' to describe what is measured by an adjutative scheme. We follow that terminology here.

Another example of a cumulative scheme is that used by the Victorian Squash federation to rank its players in pennant competition. Players are awarded rating points according to the position at which the match was played and the result of the match. The rating points allocated to the players at the four different positions within each team are outlined below:

<table>
<thead>
<tr>
<th>Position Played</th>
<th>Maximum Points</th>
<th>Match Outcome</th>
<th>Maximum Allocation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>29 Points</td>
<td>3/0</td>
<td>100%</td>
</tr>
<tr>
<td>2</td>
<td>20 Points</td>
<td>3/1</td>
<td>90%</td>
</tr>
<tr>
<td>3</td>
<td>14 Points</td>
<td>3/2</td>
<td>70%</td>
</tr>
<tr>
<td>4</td>
<td>10 Points</td>
<td>2/3</td>
<td>30%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1/3</td>
<td>10%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0/3</td>
<td>0%</td>
</tr>
</tbody>
</table>

For example, a player who wins 3 games to 1 at number 2 will earn 90% x 20 = 18 points while his opponent earns 10% x 20 = 2 points.

While such a system might give a reasonable measure of a player's performance over the year, there are several 'problems'.

- The current rating points are being applied to each individual grade in the competition, and therefore only players within the same grade can be compared.

- Because the system is cumulative, players who play more will receive more points. Of course this may be what the organisers wish to do. However a player who plays half the matches, and wins them all will be ranked no higher than a player who plays all the matches and wins half. Similarly, a good player in a lower section who is called on for half the matches to fill in at a higher grade will be ranked low in both grades.

- In common with most adjutative schemes the scheme also suffers in that a good effort can go unrewarded. Unlike the ATP tennis ranking, a player can earn points for a loss, but an extremely close 3 set loss still earns zero points.
• This rating table has a very strong tendency to cluster players together according to the position at which they are playing. For example, players performing at position one would tend to be rated all together at the top of the table, whilst players performing at position four would tend to be clustered together at the bottom of the rating table. In practice we would expect player's abilities to be spread more uniformly.

• The points are awarded on the basis of position, not quality of opponent. Similar wins (or losses) against the same opponent are given more points when that match is played at a higher position.

• The association has a lot of difficulty incorporating tournament results into the rating. Good players, absent from the competition to play tournaments, fall down the rating because they are not accumulating points. They do not receive any recognition for good performances in tournaments.

Some of these points could be rectified by making adjustments to the scheme. However such adjustments usually tend to be arbitrary and not based on any underlying theory. A Table Tennis Association has a ranking scheme which takes 4 pages to describe, that incorporates standard points (ranging from 1 to 12 depending on the difference between the opponents ratings), match weighting factors (a multiplication factor ranging from 1 to 5 in steps of .5 depending on the importance of the match), tournament bonus points (ranging from 0 to 80 depending on weighting and how far a player progresses), pennant bonus points (depending on grade), round robin bonus points (depending on the number in the tournament and the finishing position), negative points (from 5 to 15, and only deducted from the loser if they are ranked well above the winner) and dummy players (players not ranked, but given a dummy ranking by the organisers).

However all cumulative schemes essentially measure a players overall performance in the competition. They are not useful for prediction, and do not measure ability. There is a better (and simpler) way.

3. **Exponential Smoothing System**

The exponential smoothing method overcomes most of these problems. Players ratings are adjusted depending on their expected performance based on their current ratings. The only data that needs to be stored is the player's single rating.

The system works on the margin between the two players. In football the margin might be the difference in points, in squash the difference in games (we will discuss this later). For the moment suppose we have some way of determining a margin that reflects the closeness of the match. Suppose player A is rated 200, player B rated 170. Then we expect A to beat B by 30. If in fact player A beats B by 50, that implies either A is rated too low or B too high. The difference between the actual margin and the expected margin is 20, so we adjust each player's rating by some percentage (say 20%) of this. So Player A goes up by 4 to 204 and player B goes down by 4 to 166. The basic formula is
New Rating of Player

= Old rating +0.2 *[actual margin - predicted margin]
= Old rating +0.2 *[\(\text{player score} - \text{opponent score}\)]
- \(\text{player rating} - \text{opponent rating}\)]

Again in the above if A loses by 10 then we have:
New rating of A = 200 + 0.2*[\(-10 - 30\)] = 192.
New rating of B = 170 + 0.2*[\(10 - (-30)\)] = 178.

Such a system is used by Clarke [3], [4] to rate and predict football. He shows it is just as accurate as expert tipsters. Strauss & Arnold [5] suggest its use for racquetball. In chess it is called the ELO system and has been used for many years to rate all chess players, and compare players from different eras. Elo [6] gives a lot of the theory underlying the system but the book is rather difficult to obtain. The mean and the range of the ratings can be chosen arbitrarily at the expense of slight complications in the formula. There are some problems in operating such a system, particularly to preserve its integrity in rating players over a long period of time. For example the treatment of additions through new players and deletion through retirement or disinterest has to be handled carefully. However such a system has many advantages.

- It is completely objective. There is no subjective input required from organisers at all. For example, you do not have to decide how position 1 compares with position 2 - in general players with higher ratings will play in position 1. Administrators do not need to decide the importance of tournaments. Tournaments are treated the same way as any other match - the more important the tournament then the higher will be the ratings of the participants.

- The system has an underlying theoretical base. A given difference in the rating of any two players translates to an expected margin, and a probability of winning. Administrators have an idea of the difference in match performance between any 2 players. Players know how much a given rating improvement will affect their court performance.

- The system needs to store only the rating of players and so is suitable for handling a large number of players. Compare this with the ATP tennis ranking which requires the storage of the last 12 months’ results.

- The system makes every match important - and every game important. A player faced with a certain loss will still strive for every possible game. Superior players will not just take it easy if winning comfortably, as the margin of victory will affect their rating. This should improve the overall standard of the game.

4. Application to Tennis or Squash

For squash and tennis we need to investigate how 'performance' is measured - by sets, games, points etc. For squash, games seems too crude, so we need to incorporate points. This technique must ensure that the player actually winning the
match is always allocated more performance points, even in the situation where the losing player actually wins more points than the winning player (i.e. Player A beating Player B 0/9 0/9 10/9 10/9). I will use SPARKS (For Set - Point mARKS), but something like 'performance points' could be used. 20 sparks are allocated for winning a game and 1 spark is allocated for winning a point.

Case 1: A wipeout in 3 Sets (Player A beats Player B 9/0 9/0 9/0)
Player A Score: (3 Games x 20) + (27 Points x 1) = 87 sparks
Player B Score: (0 Games x 20) + (0 Points x 1) = 0 sparks
so margin for rating purposes is 87

Case 2: A Very Close 3 Setter (Player A beats Player B 10/9 10/9 10/9)
Player A Score: (3 Games x 20) + (30 Points x 1) = 90 sparks
Player B Score: (0 Games x 20) + (27 Points x 1) = 27 sparks
so margin for rating purposes is 63

Case 3: A one sided 4 Setter (Player A beats Player B 9/10 9/0 9/0 9/0)
Player A Score = (3 Games x 20) + (36 Points x 1) = 96 sparks
Player B Score = (1 Games x 20) + (10 Points x 1) = 30 sparks
so margin for rating purposes is 66.

Case 4: A Close 4 Setter (Player A beats Player B 9/5 6/9 10/9 9/7)
Player A Score = (3 Games x 20) + (34 Points x 1) = 94 sparks
Player B Score = (1 Game x 20) + (30 Points x 1) = 50 sparks
so margin for rating purposes is 44

Case 5: Closest possible 5 Set Win (Player A beats Player B 0/9 0/9 10/9 10/9)
Player A Score = (3 Games x 20) + (30 Points x 1) = 90 sparks
Player B Score = (2 Games x 20) + (45 Points x 1) = 85 sparks
so margin for rating purposes is 5

Since the maximum margin is 87 sparks the rating formula would need to be adjusted slightly to allow for this. If in going from a beginner to a world champion we have (say) 10 levels where a higher level player defeats a lower level player about 9-0, 9-0, 9-0, this implies a range of ratings about 900.

For tennis, most matches are of only 3 sets, and usually games are recorded but not points. While a more accurate measure may be obtained by including points, there is no use having a system that uses data not normally available. A possible performance measure for tennis could be 6 for each set and one for each game. Thus

Case 1: 6-0, 6-0 victory gives 2 sets 12 games to 0 sets 0 games for 24 sparks to 0
Case 2: 6-4, 6-4 victory gives 2 sets 12 games to 0 sets 8 games for 24 sparks to 8
Case 3: 6-5, 5-6, 6-5 gives 2 sets, 17 games to 1 set 16 games for 29 sparks to 22.
Case 4: 6-5, 0-6, 6-5 to give 2 sets, 12 games to 1 set, 16 games for 24 sparks to 22.
Note – the system doesn't need the actual scores, only the total number of sets, games etc. ie 2 sets to 1, 15 games to 12. Many squash and tennis score sheets require the totals to be entered in this manner anyway.

5. **Some Minor Problems**

**English and American Scoring in squash**

In English scoring, games are played to 9 points and players only win a point when they win a rally in which they serve. Winning their opponent's serve only gives them the right to serve the next rally. Thus a player who loses 9-0 may have won several rallies. In American scoring a game is up to 15 and every rally won scores a point. Thus a 9-3 win in English may be equivalent to 15-9 in American. If this complication needed to be accounted for there are several papers that address the relationship between English and American scores. eg Clarke & Norman [7], Strauss & Arnold [5], Goldstein [8].

**Doubles**

A lot of club tennis is played as doubles consisting of only one set. A player would need a separate doubles rating, and the rating of the pair determined as the average rating. Each individual player would then be adjusted in the same manner as above. Hoffman [2] gives details of using an adjustment scheme to rate individuals in a team competition. For one set matches, some adjustment to the predicted margin would be necessary.

**Testing**

Naturally such a system would need to be tested and tuned. There are various parameters that can be altered (the smoothing constant, the relative importance of sets, games and points etc). In a project for The Victorian Squash Association some testing of rating systems was performed by Clarke, Bucci et al [9]. There appears to be two ways of testing.

- Use simulation to see how well various rating schemes optimise the correlation between the ratings produced and the known ratings incorporated in the simulation.

- From actual match results, use the rating to predict results of future matches. Clarke [3] describes how several years data was used to optimise the value of the parameters in applying the method to football.

6. **Implementation**

There are various ways such a system could be implemented. At one level it could be used by a single competition to rate their players. Thus in pennant squash (or tennis) it could be run if necessary in tandem with a cumulative scheme. The cumulative scheme could still be used to decide 'most successful player' etc, but the smoothing scheme would be invaluable in deciding tournament seedings, promotion/relegation of teams, grading of new teams or other matters.
Note that the correct data must be recorded. The system requires the date of the match (probably optional, but strictly speaking ratings should be updated in order of match date; in most competitions, round number would do just as well), the names (or unique identifiers) of the two players and the match result. While this seems a minimum requirement for any match record system, different forms of recording results may mean results are not in this form. For example the computer system for the Victorian squash results kept records on a club, grade and position basis, so it was not easy to match opponents Clarke, Bucci et al [9].

Once the rating system ranges over a number of competitions, there is a necessity for unique lifelong registration numbers. There may be a possibility of nesting, with each level (state, competition, club) handling matches between their own players. While a player's rating actually moves up and down slightly with each match, it may only be necessary to run through the matches and publish updated ratings at set intervals of (say) every 3 or 6 months.

How such a system would be funded I have not investigated. However with increasing use of computers to store and process competition results, along with the increasing use of modems and electronic means of transferring data, the means of setting up a nationwide rating scheme for tennis or squash players is available. I would certainly be willing to help any competition administrators to institute such a system. I would be particularly interested in hearing from anyone with competition data in a form suitable for testing the above system.

In attempting to interest administrators in such a system, I am aware of the words of David Strauss in a private communication on his efforts to introduce a similar scheme in the U.S.A. "We didn't get too far. I didn't persuade the national people to go for it. We also developed a tennis system but didn't get it adopted. The whole thing needs marketing skills. Good Luck". While I am sure my marketing skills are no better than David's, it seems that at a conference on Mathematics and Computers in Sport, collectively we have the technical ability and the interest in the area to begin its implementation. I would be interested in discussing the possibilities with anyone at the conference.

REFERENCES


Second Conference on:

MATHEMATICS AND COMPUTERS IN SPORT

Held at Bond University, Queensland, Australia
11th to 13th July, 1994

Sponsored jointly by
The Australian Mathematical Society
and
The Australian Sports Commission

Edited by Neville de Mestre, Associate Professor of Mathematics
School of Information Technology, Bond University.