On the (Page)Ranking of Professional Tennis Players

Nicholas Dingle[1], William Knottenbelt[2], and Demetris Spanias[2]

[1] School of Mathematics, University of Manchester, Oxford Road, M13 9PL, United Kingdom. Email: nicholas.dingle@manchester.ac.uk
[2] Department of Computing, Imperial College London, South Kensington Campus, SW7 2AZ, United Kingdom. Email: {wjk,ds406}@doc.ic.ac.uk

Abstract. We explore the relationship between official rankings of professional tennis players and rankings computed using a variant of the PageRank algorithm as proposed by Radicchi in 2011. We show Radicchi's equations follow a natural interpretation of the PageRank algorithm and present up-to-date comparisons of official rankings with PageRank-based rankings for both the Association of Tennis Professionals (ATP) and Women's Tennis Association (WTA) tours. For top-ranked players these two rankings are broadly in line; however, there is wide variation in the tail which leads us to question the degree to which the official ranking mechanism reflects true player ability. For a 390-day sample of recent tennis matches, PageRank-based rankings are found to be better predictors of match outcome than the official rankings.

1 Introduction

The rankings of the world's top tennis players are the subject of much global popular interest. Indeed, a number one ranking can bring with it a great deal of prestige and celebrity, as evidenced by Association of Tennis Professionals (ATP) player Novak Djokovic's recent appearance in Time magazine's 2012 list of the Top 100 most influential people in the world[1]. Rankings can also cause a great deal of controversy, as evidenced by the recent debate over Women's Tennis Association (WTA) player Caroline Wozniacki's ranking. Wozniacki held the number one position in the WTA rankings for 67 weeks leading up to 23 January 2012, despite her failure to win a Grand Slam tournament. This led former Wimbledon champion Martina Navratilova to observe in early 2012: "If we still had the same ranking system we were using six years ago... Kvitova would have ended up number one... [Wozniacki]'s number one because that's how they set up the computer ranking... *It weighs too much on quantity and not enough on quality...* Caroline doesn't need to explain why she was number one, it's the WTA that needs to explain that."[2]

[1] See http://www.time.com/time/specials/packages/article/0,28804,2111975_2111976_2111961,00.html
[2] See http://www.reuters.com/article/2012/01/23/us-tennis-open-navratilova-idUSTRE80M0JY20120123

To some, the notion of an overall ranking might seem simplistic in a sport like tennis which features an unknown degree of transitivity (one of the primary requirements for any total ordering) and a plethora of variables that might potentially affect the outcome of any individual match – e.g. player handedness, player height, playing surface, match location, weather conditions, and even recent changes in marital status [6]. To others, the concept of a ranking system for elite sportspersons invokes moral revulsion and a questioning of the values they might be understood to promote [9]. Despite these issues, there is no denying the public's fascination with rankings, as well as the public's strong desire that any ranking system should be "fair" in some sense.

Naturally, not all interest in rankings can be ascribed to purely altruistic motives. Researchers and speculators have been keen to assess the predictive power of rankings. For example, Clarke and Dyte proposed an approach based on logistic regression to use ATP rating points to predict the outcome of tournaments [4]. Corral proposed a probit model to assess the degree to which differences in official rankings are good predictors of the outcome of Grand Slam matches [5].

Some authors have made the case that the official tennis ranking system does not actually rank players according to their relative abilities but rather simply measures their cumulative progress through various tournament rounds. This is because, under both the official ATP and WTA ranking systems (both of which will be described more fully later), points are awarded according to the highest round reached in each tournament, irrespective of the quality of opposition defeated or the margins of victory. By way of example, winning an ATP Tour 500 tournament like Memphis will yield more ranking points than making the quarter finals of a Grand Slam[3]. Nevertheless, there is a limit on the total number of tournaments that can count towards the rankings for any given player, together with the requirement to include the Grand Slams and certain more prestigious mandatory tournaments. This means that players (especially those in the top 20) cannot gain a high ranking by accumulating a large number of victories in minor tournaments alone.

Clarke [3] proposed an alternative ranking system whereby players are assigned a numerical rating which is adjusted using exponential smoothing according to the difference between the match result one would expect given the difference in ratings between the players participating and the actual match result. This difference might be measured in terms of "sparks" – or Set-Point mARKS – which are earned for winning games and points. More recently, Radicchi [8] proposed a method similar to Google PageRank [2] to rank players according to the opponents they have defeated over a period of time. The particular context was an investigation into the greatest male tennis player of all time[4].

In the present paper we compare and contrast the PageRank-style tennis rankings proposed by Radicchi with the rankings systems used in the sport today. We show an up-to-date (April 2012) comparison of official and PageRank-based

[3] Source: `http://grandslamgal.com/atp-mens-tennis-rankings-explained/`

[4] Surprisingly this turned out to be not Roger Federer or Rafael Nadal but Jimmy Connors.

rankings for both the ATP and WTA tours. We also investigate the predictive power of the official and PageRank-based ranking systems in forecasting the outcomes of matches.

The remainder of this paper is organised as follows. Section 2 describes the current ATP and WTA rankings systems. Section 3 describes the PageRank algorithm and how it can be applied to the ranking of tennis players. Section 4 presents results while Section 5 concludes.

2 Current Ranking Systems

2.1 ATP

Table 1. ATP ranking points structure for larger tournaments (excludes Challengers and Futures tournaments, the Olympics and Tour Finals)

	W	F	SF	QF	R16	R32	R64	R128	Qual.[5]
Grand Slams	2000	1200	720	360	180	90	45	10	25
Masters 1000	1000	600	360	180	90	45	10(25)	(10)	25
ATP Tour 500	500	300	180	90	45	20	-	-	20
ATP Tour 250	250	150	90	45	20	(10)	-	-	12

The current ranking system used in professional men's tennis is the South African Airways ATP Rankings, developed by the ATP in 2009 with the intention of providing an "objective merit-based method . . . for determining qualification for entry and seeding in all tournaments. . . " [1]

A player's ATP Ranking is computed over the immediate past 52 weeks, and is based on the total points a player accrues in the following 19 tournaments (18 if he did not qualify for the ATP World Tour Finals):

– The four so-called Grand Slam tournaments (Australian Open, French Open, Wimbledon US Open)
– The eight mandatory ATP World Tour Masters 1000 tournaments,
– The previous Barclays ATP World Tour Finals count until the Monday following the final regular-season ATP event of the following year.
– The best four results from all ATP World Tour 500 tournaments played in the calendar year
– The best two results from all ATP World Tour 250, ATP Challenger Tour, and Futures Series tournaments count.

In those years when the Olympics are held, results from the Olympics also count towards a player's world ranking.

As shown in Table 1, points are awarded according to the round (beginning with Qualifying, and ending with the Final) in which a player is eliminated – or if they win the tournament.

[5] Points awarded for qualifying subject to adjustment depending on draw size

2.2 WTA

Table 2. WTA ranking points structure for larger tournaments (excludes ITF Circuit tournaments, the Olympics and Tour Finals)

	W	F	SF	QF	R16	R32	R64	R128	Qual.[5]
Grand Slams	2000	1400	900	500	280	160	100	5	60
Premier Mandatory	1000	700	450	250	140	80	50(5)	(5)	30
Premier 5	800	550	350	200	110	60(1)	(1)	-	30
Premier	470	320	200	120	60	40(1)	(1)	-	20
International	280	200	130	70	30	15(1)	(1)	-	16

Similarly to ATP rankings, a player's WTA ranking is computed over the immediate past 52 weeks, and is based on the total points a player accrues at a maximum of 16 tournaments. As shown in Table 2, points are awarded according to the round in which a player is eliminated or for winning the tournament. The tournaments that count towards the ranking are those that yield the highest ranking points. These must include:

- The four Grand Slam tournaments (Australian Open, French Open, Wimbledon US Open)
- Premier Mandatory tournaments (Indian Wells, Miami, Madrid, Beijing)
- The WTA Championships (Istanbul)

For top 20 players, their best two results at Premier 5 tournaments (Doha, Rome, Cincinatti, Montreal, Toronto and Tokyo) also count[6]. As for the ATP tour, in those years when the Olympics are held, results from the Olympics also count towards a player's world ranking.

3 PageRank Applied to Tennis Players

The original formulation of PageRank [2] uses a random surfer model to measure the relative importance of web-pages. The central idea is that pages which are linked to by a large number of other pages are regarded as being more important than those with fewer incoming links; a surfer clicking through links on web-pages at random is therefore more likely to land on the more important web-pages.

For a web-graph with N pages, PageRank constructs an $N \times N$ matrix R that encodes a surfer's behaviour in terms of the matrices W, D and E.

The first behaviour modelled is where a surfer randomly clicks on links on a given page to move to another page. The corresponding matrix W has elements w_{ij} given by:

[6] Source: `http://www.wtatennis.com/SEWTATour-Archive/Ranking_Stats/howitworks.pdf`

$$w_{ij} = \begin{cases} \frac{1}{deg(i)} & \text{if there is a link from page } i \text{ to page } j \\ 0 & \text{otherwise} \end{cases}$$

where $deg(i)$ denotes the total number of links out of page i.

The second behaviour is that when a surfer encounters a page that has no outgoing links, they randomly jump to any other page in the web-graph. This is described by the matrix $D = du^T$, where d and u are column vectors:

$$d_i = \begin{cases} 1 \text{ if } deg(i) = 0 \\ 0 \text{ otherwise} \end{cases}$$

$$u_i = 1/N \quad \forall i, 1 \leq i \leq N$$

We note that other probability distributions for u are possible; here we consider only a uniformly distributed choice. The two behaviours are then combined into a single-step transition matrix $W' = W + D$.

The third and final behaviour to be modelled is that of a surfer deciding to ignore the links on the current page and to surf instead to some other random page. This is captured in a dense matrix E with elements $e_{ij} = u_j \ \ \forall i, j$.

The surfer's overall behaviour is determined by the whether or not they choose to follow the link structure of the web-graph or to jump about at random. The balance between the two is controlled by the parameter α ($0 \leq \alpha \leq 1$). The overall one-step PageRank DTMC transition matrix R is therefore defined as:

$$R = (1 - \alpha)W' + \alpha E \tag{1}$$

which is a dense matrix due to the presence of E. The PageRank of the web-graph is calculated by solving the DTMC steady-state problem:

$$x = xR \tag{2}$$

To avoid calculations with a dense matrix we rewrite Eq. 2 using Eq. 1, the definition of the matrix E and the fact that $\sum_i x_i = 1 \ \forall i$, so that the calculation only involves sparse matrices [7]:

$$x = (1 - \alpha)xW' + \alpha u$$

This can easily be manipulated into the form $Ax = b$:

$$(I - (1 - \alpha)W')^T x^T = \alpha u^T$$

and then solved for x using a method such as Conjugate Gradient Squared.

The idea of applying a PageRank-like algorithm to tennis players was first proposed by Radicchi [8]. Radicchi's formulation of the problem is equivalent to the matrix-based description of PageRank given above. When using PageRank to model tennis, the pages in the web-graph become the records of the players in their head-to-head encounters, and instead of N pages we have N players. The major difference from standard PageRank is that PageRank disregards multiple outgoing links from a single source page to a given target page, while we count the number of times a single player loses to each of their opponents.

Each player (node) in the network is assigned a "prestige score" which is passed on to other players through weighted edges. The prestige scores, P_i in a network of N nodes, can be found by solving the system of equations:

$$P_i = (1-\alpha)\sum_j P_j \frac{w_{ji}}{s_j^{out}} + \frac{\alpha}{N} + \frac{(1-\alpha)}{N}\sum_j P_j \delta(s_j^{out}) \tag{3}$$

for $i = 1, ..., N$ with the constraint $\sum_i (P_i) = 1$. In this equation, w_{ji} is the outgoing weight from player j to player i and by that we mean the number of defeats player j has suffered against player i, s_j^{out} is the total out-strength of player j (i.e. $s_j^{out} = \sum_i w_{ji}$), α is a damping parameter where $0 \leq \alpha \leq 1$ and N is the total number of players in the network. The δ function takes a value of 1 for an input of 0 and a value of 0 otherwise.

Radicchi's model defines the (i,j)th entry of W, denoted w_{ij}, as the number of matches player i has lost to player j normalised over the total number of matches player i has lost. Just as web-pages linked to by a large number of other pages will achieve a high PageRank score, so too will players who defeat a large number of other players.

The definitions of D, E and α are unchanged but we interpret them differently. We need D in the cases where a player has no defeats recorded against them – in reality this is unlikely to occur, but it may be the case in our data-sets given that we only have access to results from a limited time period. In this case, we assume the player is equally likely to lose to all other players given the absence of any information to the contrary.

Just as a surfer may disregard the links on a current page and surf to a random page, we believe that it is possible for any player to lose to any other (due to a variety of unpredictable external factors) and this is how we interpret E. The scalar parameter α lets us decide how likely we think it is that this will happen. In the experiments that follow we set α to 0.00001.

4 Results

4.1 The January 2012 WTA rankings

We return briefly to the January 2012 debate over the WTA rankings mentioned in the introduction. Table 4.1 presents a comparison of the official WTA rankings and PageRank-based WTA rankings as at 12 January 2012. While both rankings feature the same set of players in the top 10 (all of whom of are undoubtedly among the sport's elite female players), the PageRank-based rankings do appear to support the contention that Petra Kvitova may have been a more appropriate number 1 in early 2012.

Table 3. Official and PageRank-based WTA Rankings on 12 January 2012

	Official WTA Rankings		PageRank-based WTA Rankings	
1	Caroline Wozniacki	(DEN)	Petra Kvitova	(CZE)
2	Petra Kvitova	(CZE)	Vera Zvonareva	(RUS)
3	Victoria Azarenka	(BLR)	Caroline Wozniacki	(DEN)
4	Maria Sharapova	(RUS)	Victoria Azarenka	(BLR)
5	Samantha Stosur	(AUS)	Samantha Stosur	(AUS)
6	Na Li	(CHN)	Marion Bartoli	(FRA)
7	Vera Zvonareva	(RUS)	Na Li	(CHN)
8	Agnieszka Radwanska	(POL)	Agnieszka Radwanska	(POL)
9	Marion Bartoli	(FRA)	Maria Sharapova	(RUS)
10	Andrea Petkovic	(GER)	Andrea Petkovic	(GER)

4.2 Official and PageRank-based Rankings of Contemporary Players

Fig. 1 compares the ranks generated by the PageRank approach described in Section 3 with the current ATP ranks for the top 120 male players. Players located on the dashed line have the same PageRank as ATP rank; those players above the line have a higher PageRank than ATP rank while the opposite holds for those below it. We observe that the top 8 players have the same ranks under both systems, but that there is an increasing disparity between the two ranking systems for lower ranked players.

Fig. 2 compares the ranks predicted by the PageRank approach described in Section 3 with the current WTA ranks for the top 120 female players. Again, we observe that agreement between the two ranking systems is best for higher ranked players, although even within the higher ranked players there are some surprising differences. This might be because of the gentler (relative to the ATP rankings) gradient between the score achieved by a tournament winner compared to players reaching later tournament rounds.

The seeding system used in tournaments may explain why there is less agreement between the PageRank and official ranks for the weaker players. Lower ranked players are more likely to be matched with higher ranked players in the initial rounds, and this makes it harder for the weaker players to proceed. This has two possible effects on rankings. First, weaker players have less opportunity to proceed to the later rounds of tournaments where the ranking points received per victory are significantly higher. In contrast, under PageRank players are compensated with an appropriate amount of PageRank when they defeat an opponent of a given level of ability irrespective of the round. Second, lower ranked players tend to play fewer tournament games than high ranked ones and this limits the amount of data on which to base rankings under any system.

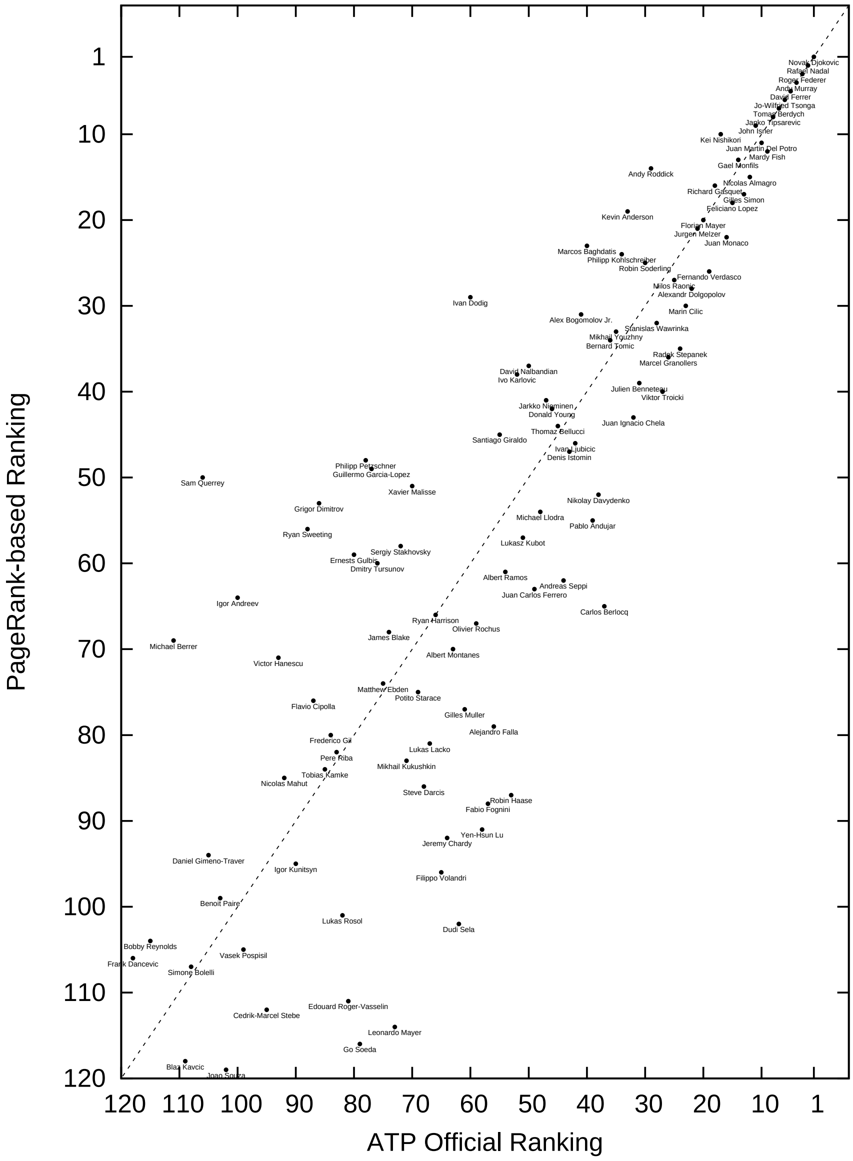

Fig. 1. PageRank-based Ranking vs. Official Ranking for ATP players (April 2012)

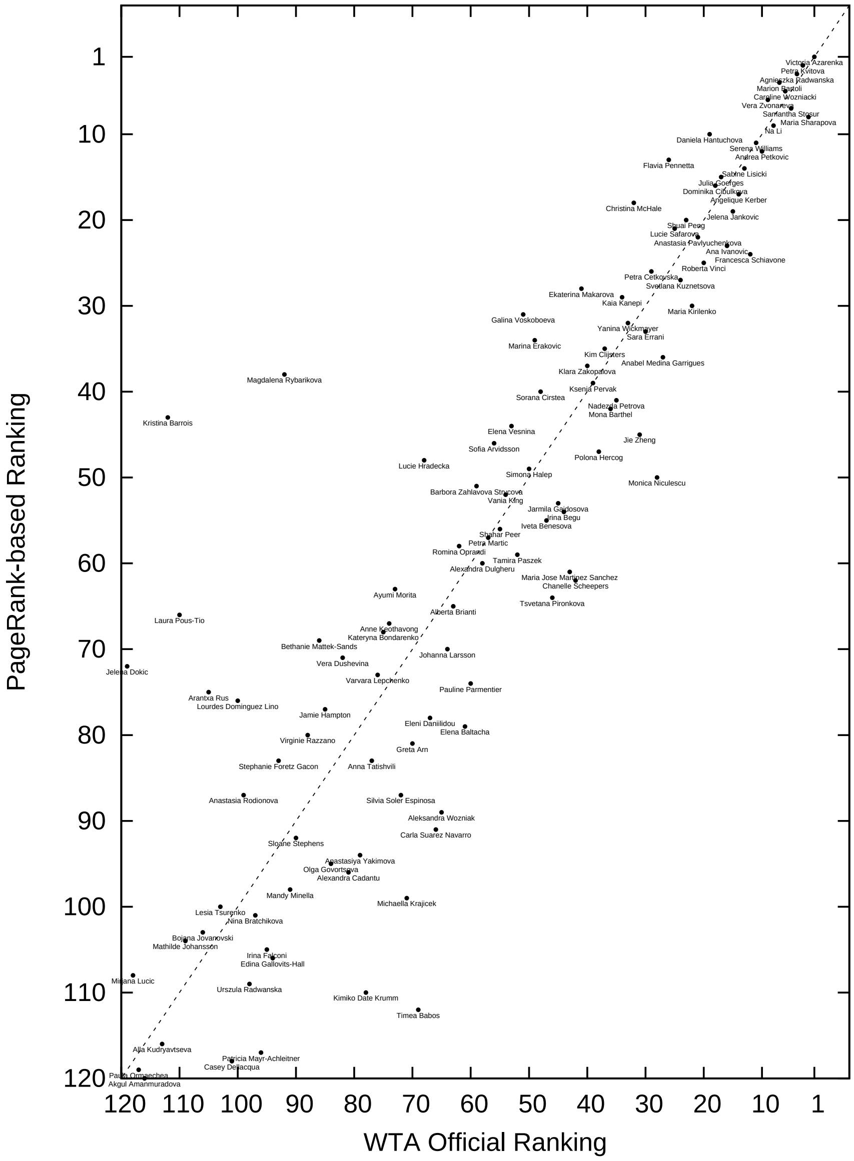

Fig. 2. PageRank-based Ranking vs. Official Ranking for WTA players (April 2012)

4.3 Predictive Power

We are interested in comparing the predictive power of ranking systems. The simplest approach to forecasting the winner of a tennis match is to select the player with the lowest rank. We now investigate how good the official system and the PageRank approach are when used in this way.

Table 4. Predictive power of official and PageRank tennis rankings over 390 days of recent matches.

Tour	# Matches	Official			PageRank		
		Right	Wrong	%	Right	Wrong	%
ATP	12 022	7 987	4 035	66.4%	8 055	3 967	67.0%
WTA	12 775	8 406	4 369	65.8%	8 470	4 305	66.3%

Table 4 compares the success of using the two approaches over 390 days worth of historical matches played in 2011/2012. We observe that approximately 66% of the time selecting the lower ranked player is correct, and that this percentage is about the same regardless of how the player's rank is computed.

Table 5. Predictive power of official and PageRank tennis rankings in those cases where the different systems predict different outcomes.

Tour	# Matches	Official			PageRank		
		Right	Wrong	%	Right	Wrong	%
ATP	1 738	835	903	48.0%	903	835	52.0%
WTA	1 876	906	970	48.3%	970	906	51.7%

There are times, however, where the prediction based on official rank differs from that produced by PageRank. In this situation, as shown by the results in Table 5, there is an advantage to using the PageRank results over the official rankings because they predict the correct outcome correctly more often (approximately 52% of the time as opposed to 48% with the official ranks).

5 Conclusions and Future Work

We have taken Radicchi's PageRank-inspired tennis ranking system and applied it to calculate rankings for players currently playing on the ATP and WTA tours. We observed that the two systems tend to rank the top players consistently but that there is considerable disagreement for lower-ranked players. We believe this can be attributed to the seeding system used in tournaments. We have also investigated the use of the two ranking approaches to predict the outcome of tennis matches and have observed that, when the predictions so generated differ, the PageRank approach appears to be (on average) a slightly better predictor.

In future we could conduct a wider experiment similar to [5], which investigated whether differences in rankings were good predictors of the outcome of Grand Slam tennis matches, but using PageRank-based ranking rather than official rankings. This will build on the prediction work reported in this paper. It may be interesting to experiment with PageRank-based systems that take into account the margin of victory of matches and to see if this approach yields greater predictive power. We could also investigate the predictive power of a more fine-grained PageRank-based approach that is constructed from set-level results, rather than the match-level results presented in this paper.

We would like to investigate the parameter α in more detail, both to assess the sensitivity of our results to particular values and to estimate accurate values from available statistics. We will also evaluate to what extent the length of the match (3 sets or 5 sets) has on the chances of an upset, and whether this could account for the fact that our PageRank-derived rankings more closely match the official rankings for male players than for female players.

Acknowledgements

The authors would like to thank the anonymous referee for their insightful comments and suggestions.

References

1. ATP Tour, Inc: The 2009 ATP Official Rulebook (2011), http://www.ATPWorldTour.com
2. Brin, S., Page, L.: The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems 30(1-7), 107–117 (1998), http://www.sciencedirect.com/science/article/pii/S016975529800110X, proceedings of the Seventh International World Wide Web Conference
3. Clarke, S.R.: An adjustive rating system for tennis and squash players. In: de Mestre, N. (ed.) 2nd Conference on Mathematics and Computers in Sport. pp. 43–50 (1994)
4. Clarke, S.R., Dyte, D.: Using official ratings to simulate major tennis tournaments. International Transactions in Operational Research 7(6), 585 (2000)
5. del Corral, J., Prieto-Rodriguez, J.: Are differences in ranks good predictors for Grand Slam tennis matches? International Journal of Forecasting 26(3), 551–563 (2010)
6. Farrelly, D., Nettle, D.: Marriage affects competitive performance in male tennis players. Journal of Evolutionary Psychology 5, 141–148 (2007)
7. Langville, A., Meyer, C.: Deeper inside PageRank. Internet Mathematics 1(3), 335–380 (2004)
8. Radicchi, F.: Who is the best player ever? A complex network analysis of the history of professional tennis. PLoS ONE 6(2), e17249 (02 2011)
9. Tännsjö, T.: Chapter 1: Is it fascitoid to admire sports heroes? In: Tännsjö, T., Tamburrini, C. (eds.) Values in Sport: Elitism, Nationalism, Gender Equality and the Scientific Manufacturing of Winners. pp. 9–23 (2000)