Computers and Mathematics with Applications 64 (2012) 3820–3827

Contents lists available at SciVerse ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

A common-opponent stochastic model for predicting the outcome of professional tennis matches

William J. Knottenbelt, Demetris Spanias *, Agnieszka M. Madurska

Department of Computing, Imperial College London, South Kensington Campus, London, SW7 2AZ, United Kingdom

ARTICLE INFO

Keywords:
Stochastic modelling
Tennis
Sport

ABSTRACT

Tennis features among the most popular sports internationally, with professional matches played for 11 months of the year around the globe. The rise of the internet has stimulated a dramatic increase in tennis-related financial activity, much of which depends on quantitative models. This paper presents a hierarchical Markov model which yields a pre-play estimate of the probability of each player winning a professional singles tennis match. Crucially, the model provides a fair basis of comparison between players by analysing match statistics for opponents that both players have encountered in the past. Subsequently the model exploits elements of transitivity to compute the probability of each player winning a point on their serve, and hence the match. When evaluated using a data set of historical match statistics and bookmakers odds, the model yields a 3.8% return on investment over 2173 ATP matches played on a variety of surfaces during 2011.

1. Introduction

Tennis has been growing in popularity around the globe ever since the first Wimbledon Championships were held at the All England Club in 1877, attracting an audience of around 200. Indeed, in China alone, an estimated audience of 65 million[1] watched Li Na play Francesca Schiavone in the 2011 French Open Women's final. Stimulated by the growth of the internet, this rise in popularity has been accompanied by a dramatic increase in the financial activity related to tennis, both in terms of traditional bookmaking and modern betting exchange volumes. For example, on the Betfair betting exchange more than £40 million was traded in the Match Odds market of the 2011 US Open Men's final. In recent times hedge funds have been set up to exploit the excellent risk diversification opportunities offered by sporting events including tennis matches.[2] A significant proportion of tennis-related financial activity depends on the availability of mathematical models for predicting the probability of victory of each player. It is such a model which forms the focus of this paper.

Professional singles tennis is an attractive sport to model mathematically for a number of reasons. Firstly, each match is played between just two players, as opposed to the multitude of players involved in a team-based game such as football, rugby or basketball. Because of this there is no need to analyse the offensive and defensive strengths of possible combinations of starting line ups; nor do we need to consider the likelihood and potential effects of events such as substitutions or players being sent off. Secondly, a wealth of statistics about every match are gathered and made publically available both by professional bodies and web sites. The detail of these statistics is bound to increase in the future due to the use of systems such as Hawk-Eye, and perhaps future innovations in video and audio processing of match footage such as the rally analysis proposed by Hunter [1]. This facilitates model parameterisation using historical data and backtesting. Thirdly, there are only

* Corresponding author. Tel.: +44 7588053479.
E-mail addresses: wjk@doc.ic.ac.uk (W.J. Knottenbelt), ds406@doc.ic.ac.uk (D. Spanias), amm208@doc.ic.ac.uk (A.M. Madurska).

1 http://www.guardian.co.uk/sport/2011/jun/04/french-open-2011-li-na-francesca-schiavone.

2 http://www.businessweek.com/magazine/content/10_29/b4187069936116.htm.

0898-1221/$ – see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.camwa.2012.03.005

two possible outcomes of a match, whereas in sports such as horse racing numerous outcomes are possible. Finally, scoring in tennis is fine-grained resulting in more gradual in-play odds movements.

There have been numerous attempts to solve the problem of modelling a tennis match. Clarke and Dyte, Klaassen and Magnus, and Radicci approached tennis modelling using player rankings [2–4]. Liu, Barnett, Newton and Keller, and O'Malley all generated equivalent hierarchical expressions to model the probabilities of winning a tennis game, set and match based on the probability of a player winning a point while serving [5–8].

Tennis is an ideal candidate for a hierarchical model as a match consists of a sequence of sets, which in turn consist of a sequence of games, which in turn consist of a sequence of points. By making the assumption that points are independent and identically distributed (i.i.d.) [9], one can create a hierarchical model for the match which only depends on the probabilities of the two players winning a point while serving.

This is why authors like Barnett and Clarke, Newton and Aslam, and Spanias and Knottenbelt focus in detail on the estimation of that variable [10–12].

In particular, the technique presented by Barnett and Clarke, subsequently adopted by Newton and Aslam, and Spanias and Knottenbelt, approaches this problem by averaging each player's match statistics over a period of time (optionally filtered by surface type). This represents the performance of a player against the *average* opponent faced during that time period rather than a *particular* opponent.

Barnett and Clarke attempt to combine the serving ability of one player against an average opponent with the returning ability of the other player against an average opponent. To do this, their approach computes two statistics for each player, f_i and g_i, reflecting their serving and receiving strengths against an average opponent respectively [10]:

$$f_i = a_i b_i + (1 - a_i)c_i$$
$$g_i = a_{av} d_i + (1 - a_{av})e_i$$

where:

f_i = proportion of points won on serve by player i,
g_i = proportion of points won on return by player i,
a_i = probability of successful first serve by player i,
a_{av} = average probability of successful first serve (across all players),
b_i = proportion of own successful first serves won by player i
c_i = proportion of own second serves won by player i
d_i =proportion of first serves of opponent won on return by player i, and
e_i = proportion of second serves of opponent won on return by player i.

Now for the particular case where player i plays a match against player j the proportion of points won on serve by player i, f_{ij}, is estimated as [10]:

$$f_{ij} = f_t + (f_i - f_{av}) - (g_j - g_{av}) \tag{1}$$

where:

f_t = average percentage of points won on serve for tournament,
f_{av} = average percentage of points won on serve (across all players), and
g_{av} =average percentage of points won on return (across all players).

Although intuitively appealing, this solution is not perfect. This is because a good player is more likely to advance in a given tournament and face other strong players. At the same time a weaker player will usually drop out earlier, and on average play against less demanding opponents. This immediately distorts any results, as for each of the two players the notion of the "average opponent" can be quite different.

The method we propose here eliminates this bias by exploiting statistics from matches played against common opponents (i.e. opponents that both players faced in the past). The intuition behind our approach is that tennis is a transitive sport to a certain extent. Thus, we should be able to exploit the fact that if Player A is better than Player C and Player C is better than Player B, then (usually) Player A is better than Player B. The challenge is to combine the results of matches between A and C and the results of matches between B and C to reason about a match between A and B.

In the case of professional singles tennis matches, this reasoning has the potential to be especially fruitful, because of the limited number of players involved. There are roughly only 150 active professional tennis players in each of the ATP and WTA tours. Within each tour, players frequently play against each other in a variety of tournaments. Although there are a limited number of head-to-head encounters for any two given players, many pairs of players share a rich set of common opponents.

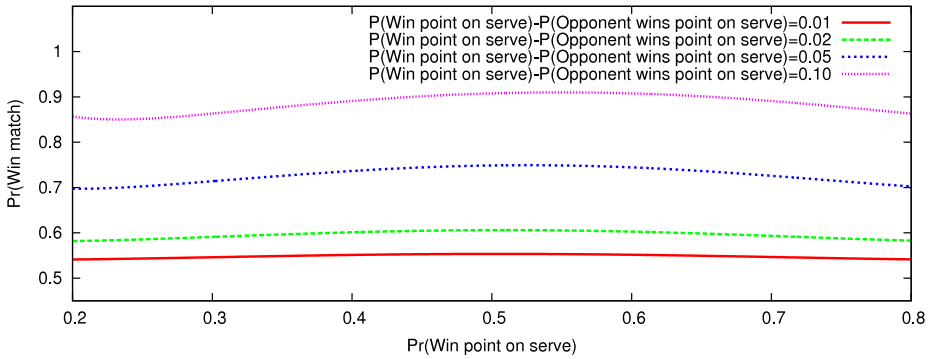

Fig. 1. Probability of the better player winning a best-of-three-sets tennis match with fixed differences of 0.01, 0.02, 0.05 and 0.10 in the two players' probability of winning a point on serve.

2. O'Malley's tennis formulae

When calculating our results we used the hierarchical equations developed by Barnett [13] and O'Malley in combination with our proposed method of calculating the probability of a player winning a point on serve. Both equations assume that the probabilities of a player winning a point during serve and return are constant throughout the match and are independent of their previous values. We will briefly describe the O'Malley's closed-form equations here as they will be referred to in the application of our model.

O'Malley derives the probabilities of a certain player winning a match, set, game and tiebreaker, assuming that winning any point in play is a Bernoulli random variable, arriving at the following formulae [8]:

Probability of winning a game:

$$G(p) = p^4\left(15 - 4p - \frac{10p^2}{1 - 2p(1-p)}\right).$$

Probability of winning a tie-breaker:

$$\mathrm{TB}(p, q) = \sum_{i=1}^{28} \mathrm{A}(i, 1)p^{\mathrm{A}(i,2)}(1-p)^{\mathrm{A}(i,3)}q^{\mathrm{A}(i,4)}(1-q)^{\mathrm{A}(i,5)}d(p, q)^{\mathrm{A}(i,6)}. \tag{2}$$

Probability of winning a tiebreaker set:

$$S(p, q) = \sum_{i=1}^{21} \mathrm{B}(i, 1)G(p)^{\mathrm{B}(i,2)}(1 - G(p))^{\mathrm{B}(i,3)}G(q)^{\mathrm{B}(i,4)}(1 - G(q))^{\mathrm{B}(i,5)} \times (G(p)G(q) + (G(p)(1 - G(q)) + (1 - G(p))G(q))TB(p, q))^{\mathrm{B}(i,6)}.$$

Probability of winning a best-of-three-sets match:

$$M_3(p, q) = S(p, q)^2[1 + 2(1 - S(p, q))]. \tag{3}$$

Probability of winning a best-of-five-sets match:

$$M_5(p, q) = S(p, q)^3[1 + 3(1 - S(p, q)) + 6(1 - S(p, q))^2] \tag{4}$$

where:

p = probability of the player winning a point on serve.

q = probability of the player winning a point on return. A, B are coefficient matrices defined in the Appendix of [8].

$d(p, q) = pq[1 - (p(1 - q) + (1 - p)q)]^{-1}$.

A fundamental insight arising from O'Malley's equations is that the probability of winning a match is mainly dependent on the *difference* of the probabilities of players winning a point while serving [8]. The graph in Fig. 1 demonstrates this by showing a plot of the better player's probability of winning the match for various fixed differences in the two players' probability of winning a point on serve. Note the x-axis runs from 0.2 to 0.8 corresponding to the domain of values likely to be encountered in professional tennis.

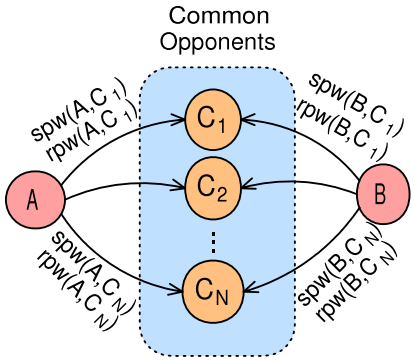

Fig. 2. Parameters of the common-opponent model.

3. Match probabilities using common-opponent model

Let players A and B be the two players playing in the match we wish to model. Also, let C_i for $1 \leq i \leq N$ be the N common opponents they have faced in the past. For each C_i we denote $spw(\mathrm{A}, C_i)$ as the proportion of service points won by A against C_i and $spw(\mathrm{B}, C_i)$ as the proportion of service points won by B against C_i. Similarly, $rpw(\mathrm{A}, C_i)$ is the proportion of returning points won by A against C_i and $rpw(\mathrm{B}, C_i)$ is the proportion of returning points won by B against C_i. This is illustrated in Fig. 2. In cases where either A or B has faced the common opponent C_i in multiple matches during the period of the data set, then $spw(\mathrm{A}, C_i)$, $rpw(\mathrm{A}, C_i)$, $spw(\mathrm{B}, C_i)$ and $rpw(\mathrm{B}, C_i)$ represent the averages over those matches.

As discussed earlier, following O'Malley's findings [8], the difference in service points won can be used as an indicative measure of the probability of a player winning the match against an opponent. In order to model how A and B would play against each other through their common opponents, C_i, we first need to calculate the differences in service points won of A and B against those opponents. Then we can additively combine those differences to come up with an indication of how well A would perform against B.

For each common opponent, C_i, we compute Δ_i^{AB} which represents a measure of the advantage (or if negative, disadvantage) Player A has over Player B in terms of the proportion of service points won against opponent C_i, as follows:

$$\Delta_i^{\mathrm{AB}} = (spw(\mathrm{A}, C_i) - (1 - rpw(\mathrm{A}, C_i))) - (spw(\mathrm{B}, C_i) - (1 - rpw(\mathrm{B}, C_i))). \tag{5}$$

This value can be used to additively influence an arbitrary probability of winning a point on serve for player A or player B in any hierarchical model. As an example, we will use O'Malley's equations to show how one can approximate the probability of Player A winning a best-of-three-sets match against Player B based on their past performances against common opponent C_i:

$$\Pr(\mathrm{A\ beats\ B\ via}\ C_i) \approx \frac{M_3(0.6 + \Delta_i^{\mathrm{AB}}, (1 - 0.6)) + M_3(0.6, (1 - (0.6 - \Delta_i^{\mathrm{AB}})))}{2}. \tag{6}$$

In Eq. (6), we calculate the match probabilities twice: once by positively influencing Player A's probability of winning a service point and once by negatively influencing Player B's probability of winning a service point. Subsequently, we average the two values. We arbitrarily use the value 0.6 because it is the approximate average probability of a professional player winning a point on serve when playing against another professional player (recall that, as shown in Fig. 1, the exact value is not critical).

To combine all the available data from all common opponents, we calculate the average of $\Pr(\mathrm{A\ beats\ B\ via}\ C_i)$ over all common opponents, C_i, to estimate the probability of player A winning the match as follows:

$$P_{\mathrm{avg}}^{\mathrm{AB}} = \frac{\sum_{i=1}^{N} \Pr(\mathrm{A\ beats\ B\ via}\ C_i)}{N}. \tag{7}$$

As an example to illustrate the above approach, consider the first round US Open 2011 match played between Vania King and Greta Arn. We base our calculations on the ten common opponents which the players previously faced on a hard surface (found by intersecting the opponents from the previous 50 matches played on a hard surface by King and by Arn).

Table 1 presents the percentage *spw* and *rpw* against the opponents for each player. We combine the above to estimate the advantage or disadvantage King has over Arn, using Eq. (5). Subsequently we calculate the probability of King winning a match with Arn, given the information inferred from each particular common opponent.

Averaging out the results presented in Table 2 using Eq. (7) gives us an estimated probability of 0.5962 of Vania King winning the match. In the event King won 6–1 6–4. By contrast, betting markets were favouring Greta Arn with an implied probability of 0.52.

Table 1
Statistical data on matches played with common opponents for Vania King and Greta Arn.

Common opponent	King		Arn	
	spw(%)	*rpw*(%)	*spw*(%)	*rpw*(%)
Alexa Glatch	60.55	42.48	67.47	39.76
Andrea Petkovic	54.93	33.75	52.86	44.95
Caroline Wozniacki	44.97	36.24	46.97	23.68
Kurumi Nara	72.97	54.69	59.68	47.37
Maria Jose Martinez Sanchez	58.33	33.33	52.38	35.14
Rika Fujiwara	56.04	55.06	53.26	47.52
Sandra Zahlavova	59.05	45.19	67.35	52.27
Sara Errani	37.29	45.00	56.14	46.55
Shuai Peng	61.64	38.36	44.00	23.68
Simona Halep	61.67	54.69	52.70	37.50

Table 2
Probability of Vania King winning against Greta Arn, given data on each of the common opponents separately.

Opponent	Probability of King beating Arn (%)
Alexa Glatch	29.38
Andrea Petkovic	12.11
Caroline Wozniacki	91.11
Kurumi Nara	99.42
Maria Jose Martinez Sanchez	70.39
Rika Fujiwara	90.62
Sandra Zahlavova	2.67
Sara Errani	0.62
Shuai Peng	99.99
Simona Halep	99.90
Overall average	59.62

One might note that there can be quite some variation between the probabilities of winning resulting from different common opponents. This suggests that predictions made with only a small number of common opponents should be treated with caution. However, our experience is that matches between active professional tennis players usually feature a sufficiently rich set of common opponents to yield a stable estimate.

This approach is not limited to O'Malley's equations and can be used with any hierarchical model which requires the probability of winning a point on serve as its input. We demonstrate this in the results section by modelling ATP matches using Barnett's hierarchical equations in combination with our estimation of the players' probabilities of winning a point on serve.

4. Results

In order to assess the efficacy of our model, we have implemented it using the match outcome formulae suggested by both Barnett and O'Malley. Subsequently, we have calculated predictions for 500 WTA matches and 2173 ATP tennis matches, based on historical statistical data from sources such as the ATP World Tour[3] and TennisInsight[4] websites. To evaluate the performance of our method, we examined whether our model was profitable when using a simple betting strategy together with the best available closing odds offered by four major bookmakers (as retrieved from the Tennis Data[5] website). Our decision to bet in the case of each match relies on the comparison of the market odds and our predicted odds for both players. For each match, we would place a £1 bet on the predicted winner if the odds that were offered by the bookmakers were better than the ones given by our model. Subsequently we calculate our winnings or losses based on the actual outcome of the match. The sum of those is treated as our absolute return over a given set of matches.

Table 3 shows how our common-opponent model combined with O'Malley's formulae performs for matches played in the 2011 WTA major tournaments. The first column shows the tournament name, the second column shows the number of matches played in that tournament (excluding retirements and walkovers) and the third column shows the number of matches that we attempted to model. Some matches have not been considered here, due to lack of common opponents in the background data. For this particular backtest we examined the last 50 matches each player played on a relevant surface,

[3] http://www.atpworldtour.com.
[4] http://www.tennisinsight.com.
[5] http://www.tennis-data.co.uk.

Table 3
Women's Tennis Association (WTA) major tournament tests using O'Malley's equations and data from players' last 50 matches played on same surface. The combined ROI on all four tournaments amounts to approximately 6.85%.

Tournament	Matches	Attempts	Success (%)	Bets (£)	ROI (%)
Australian Open	126	120	70.00	65	6.15
French Open	125	106	65.09	65	−11.62
Wimbledon	128	78	66.67	46	32.50
US Open	121	112	72.32	64	7.89

Table 4
Association of Tennis Professionals (ATP) major tournament tests using Barnett's equations and data from player's past 12 months of activity played on same surface. Combined ROI on all four tournaments amounts to approximately 6.48%.

Tournament	Matches	Attempts	Success (%)	Bets (£)	ROI (%)
Australian Open	127	89	68.54	87	6.15
French Open	126	75	69.33	75	10.85
Wimbledon	127	17	58.82	17	−23.94
US Open	127	89	77.53	86	9.02

Table 5
Association of Tennis Professionals (ATP) major tournament tests using Barnett's equations and data from player's past 12 months of activity played on all surfaces. Combined ROI on all four tournaments amounts to approximately −1.29%.

Tournament	Matches	Attempts	Success (%)	Bets (£)	ROI (%)
Australian Open	127	110	70.00	107	−3.15
French Open	126	113	75.22	112	6.43
Wimbledon	127	112	67.86	107	−11.35
US Open	127	109	75.23	106	2.60

prior to the modelled match. One could increase the number of common opponents by going further back in the history and thus have a greater number of attempted matches; however, this could negatively affect the quality of the results as it would not reflect the recent form of the players accurately. The fourth column of the table shows the percentage of successful predictions (foreseeing the actual winner) for each tournament. The fifth column is the total amount of money placed as virtual bets throughout the tournament. The return on investment is based on the profit that would result out of all the bets (not including tax or commission).

From Table 3, we can see that for individual tournaments we get fairly unstable results. This is because the sample considered, based on one individual tournament, is rather small. The combined results from all four of the major tournaments is more representative with a return on investment of approximately 6.85%. With a more sophisticated betting strategy, for example a strategy which takes into account the number of common opponents used to model the match or one that makes use of the Kelly criterion [14], it might be possible to achieve an even higher ROI.[6]

Table 4 gives an overview of the results from modelling the ATP major tournament matches played in 2011 using Barnett's equations in combination with our approach. The columns shown have been generated in the same way as they were in Table 3. It is noteworthy that similarly to the women's overall ROI, the overall ROI in Table 4 is positive and amounts to 6.48%. The data set from which common opponents were retrieved in this case consisted of matches from the past 12 months of the player's activity which were played on the same surface that the match modelled was played. This surface filtering limits the number of available data which is why there are relatively few matches attempted. This especially impacts the Wimbledon case because there are fewer tournaments on grass than on the other surfaces.

To overcome this limitation, we also generated the same table using data from all surfaces dating back 12 months from the date of the modelled match. This greatly increased the attempted matches as well as the success percentage due to the wealth of data. On the other hand, not using surface specific statistics decreased the overall ROI. These results are shown in Table 5.

Increasing our sample, we further tested our proposed approach using Barnett's model for 2173 matches played from the 1 January 2011 and 12 September 2011 in ATP tournaments. We ran this twice using different sets of statistical data. The first set of data was made up of all player activity dating back 12 months from the date the match being modelled was played. The second set was the same as the first one but excluded matches which were played on a different surface than the match being modelled.

[6] The Kelly criterion outlines a strategy for calculating the proportion of a total bankroll to wager on a bet, given the predicted probability of winning and the payoff odds provided by the bookmaker, such that the future bankroll is maximized in the long term.

Table 6
Association of Tennis Professionals (ATP) Matches played between 1 January 2011 and 12 September 2011 using Barnett's equations.

Data	Matches	Attempts	Success (%)	Bets (£)	ROI (%)
Matched surfaces	2173	1228	63.88	1204	3.80
All surfaces	2173	1873	65.46	1838	3.41

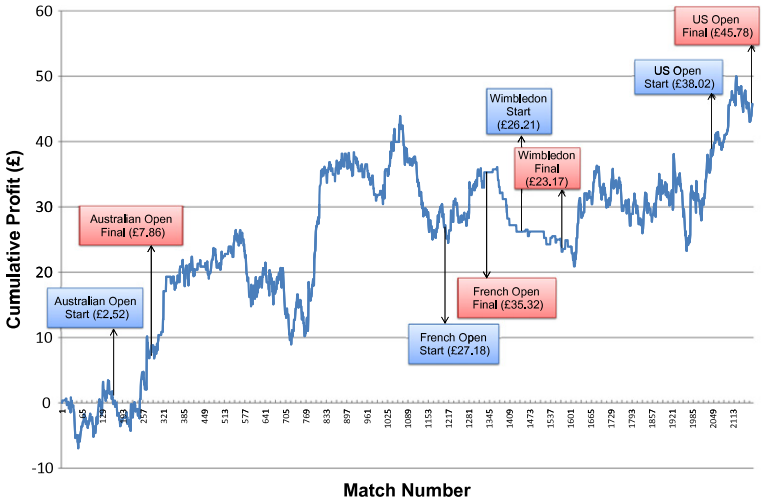

Fig. 3. Cumulative profit when using surface filter for 2173 ATP matches played during 2011.

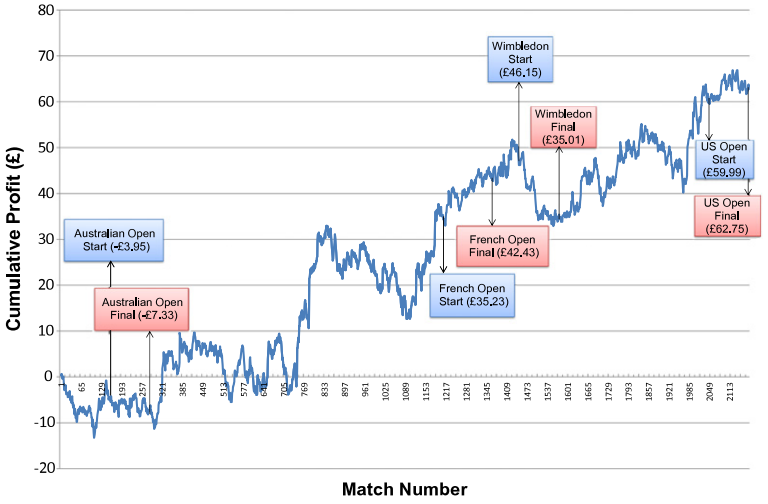

Fig. 4. Cumulative profit when using all surface data for 2173 ATP matches played during 2011.

From the results displayed in Table 6 we can see that as before, there are more attempted matches in the case of no surface filtering and at the same time there is a higher success percentage but a lower ROI. The sample of 2173 matches is big enough to represent the average performance of our model. In both cases, we get a positive return on investment which is a good indication that the model is effective over the longer term.

Figs. 3 and 4 show the cumulative profit as calculated for the data used to generate Table 6. Fig. 3 shows the cumulative profit when using statistics from matches played on the same surface as the modelled match. Fig. 4 shows the cumulative profit when using statistical data from matches played on all surfaces. The matches are sorted by the date they were played. We have marked on both graphs the beginnings and ends of some major tournaments as well as the profit values at the corresponding points. One can clearly see that these values reflect the findings shown in Tables 4 and 5.

It can be observed in both graphs that the cumulative profit has an upward trend. It is also boosted steeply upwards at particular points by a few correctly predicted upsets– i.e. results predicted as highly unlikely by the bookmakers. Even though the return on investment is slightly greater when using a surface filter, the amount of profit accumulated by the

end of the 2173 matches is greater when using data from all surfaces. This is because more matches are attempted due to a greater wealth of data, and thus a greater number of wagers are simulated. This illustrates how important the wealth of data is to our approach but also emphasises the trade-off of wealth of data against accuracy of data. Going further back in time and disregarding the surface on which a match was played on will increase the wealth of statistical data, but will also degrade the quality of the data since factors such as recent form and surface dependency are ignored.

5. Conclusion

This paper has presented a hierarchical Markov model which yields a pre-play estimate of the probability of each player winning a professional singles tennis match. The model is parameterised by analysing match statistics for opponents that both players have encountered in the past and has proved to generate 3.8% long-term profit against the best odds offered by bookmakers for a large data set of over 2000 diverse tennis matches. The bookmakers' odds we have considered were lower than the ones offered by betting exchanges such as Betfair. Thus, we conclude that our approach may have the potential to enhance returns from existing stochastic models of tennis.

However, there is most certainly room for improvement and future work. The results seem to be dependent on the number of matches in the sample investigated. The efficacy of more sophisticated betting strategies also needs to be assessed. It might also be interesting to investigate a recursive approach to the problem. In this case, we would extend the algorithm by considering common opponents between both given players and their common opponents. There is scope for research with respect to determining the optimal depth of recursion as well as an appropriate limit on the number of common opponents considered at each stage.

References

[1] G. Hunter, A. Shihab, K. Zienowicz, Modelling tennis rallies using information from both audio and video signals, in: Proceedings of the IMA International Conference on Mathematics in Sport, 2007, pp. 103–108.
[2] S.R. Clarke, D. Dyte, Using official ratings to simulate major tennis tournaments, International Transactions in Operational Research 7 (6) (2000) 585.
[3] F. Klaassen, J. Magnus, Forecasting the winner of a tennis match, European Journal of Operational Research 148 (2) (2003) 257–267.
[4] F. Radicchi, Who is the best player ever? A complex network analysis of the history of professional tennis, PLoS One 6 (2) (2011) e17249. http://dx.doi.org/10.1371/journal.pone.0017249.
[5] Y. Liu, Random walks in tennis, Missouri Journal of Mathematical Sciences 13 (3) (2001).
[6] T.J. Barnett, S.R. Clarke, Using Microsoft Excel to model a tennis match, in: G. Cohe (Ed.), 6th Australian Conference on Mathematics and Computers in Sport, 2002, pp. 63–68.
[7] P.K. Newton, J.B. Keller, Probability of winning at tennis I. Theory and data, Studies in Applied Mathematics 114 (3) (2005) 241–269.
[8] A.J. O'Malley, Probability formulas and statistical analysis in tennis, Journal of Quantitative Analysis in Sports 4 (2) (2008) 15.
[9] F.J.G.M. Klaassen, J.R. Magnus, Are points in tennis independent and identically distributed ? evidence from a dynamic binary panel data model, Journal of the American Statistical Association 96 (454) (2001) 500–509.
[10] T.J. Barnett, S.R. Clarke, Combining player statistics to predict outcomes of tennis matches, IMA Journal of Management Mathematics (2005) 113–120.
[11] P.K. Newton, K. Aslam, Monte Carlo tennis: a stochastic Markov chain model, Journal of Quantitative Analysis in Sports 4 (3) (2009) 1–42.
[12] D. Spanias, W.J. Knottenbelt, Quantitative modelling of singles and doubles tennis matches, in: Proceedings of the 3rd IMA International Conference on Mathematics in Sport, 2011.
[13] T.J. Barnett, Mathematical modelling in hierarchical games with specific reference to tennis, Ph.D. Thesis, Swinburne University of Technology, Melbourne, Australia, 2006.
[14] J.L. Kelly, A new interpretation of information rate, Bell System Technical Journal 35 (1956) 917–926.