OPERATIONS RESEARCH IN TENNIS

By

Tristan Barnett

Preface

Operations research is concerned with solving real-world problems using analytical techniques. Tennis is no exception. This was demonstrated in a book by the author 'The Mathematics of Tennis' on a match between Isner and Mahut at the 2010 Wimbledon championships that lasted more than 11 hours. This book will address solutions to solving this problem by devising alternative scoring systems to reduce the length of matches. A fairer method to the line call challenge system will also be addressed as another regulation aspect. Performance aspects obtained are the analysis of tennis data, serving strategies and resource allocation. It should be noted that predictions were analysed in 'The Mathematics of Tennis'. This book is a collection of papers written by the author and contribute to a course on tennis statistics.

Contents

Alternative Scoring Systems

1. Introduction

When analyzing tennis scoring systems it is useful to reflect on the history of tennis, as this helps in building an argument as to what are “reasonable” scoring systems for today. For example, why does a game of tennis contain four points to win followed by deuce if the scores are level after six points have been played? Would tennis be different if the next player to win the point if deuce is obtained is the winner, or an unbalanced game that required the server to win four points and the receiver to win three points without actually playing a deuce? Why does the US Open play a tiebreak deciding set, whereas the other three grand slam events use an advantage final set?

It can be important to understand changes in playing equipment, court dimensions, surface composition and player physique to help identify what scoring systems could be used for today. Parsons (2006) states “these days when there is so much concern about the dominance of power in the games, especially in men’s tennis, it is fascinating to discover that there was similar anxiety at the end of the first Wimbledon Championships in 1877. An analysis was carried out of all the scorecards and concluded from the high proportion of service games won that the service was far too powerful for the long-term good of the game”. Three remedies were suggested; heighten the middle of the net, to do way with first faults or to move the service line closer to the net. The All England Club chose the latter of reducing the size of the service court to 22ft.

The paper begins in section 2 with the history of tennis scoring systems. Section 3 looks at possible factors causing long matches to occur with examples of long matches that have occurred in recent times. Section 4 analyzes a range of scoring systems by identifying relevant characteristics, obtaining numerical results and giving recommendations of possible scoring systems that could be used today in men’s and women’s singles and doubles tournaments. Section 5 gives some concluding remarks.

2. History of tennis scoring systems

Various theories have been given for the origins of tennis. One theory is that the game was given to the French Royal Court in the 10th century by a wandering minstrel. However, by the 11th century early tennis was being played in French monasteries. Hands were used to hit the ball to begin with; gloves were used later on and during the 13th century players started to use short bats. The French called this form of tennis *jeu de paume*, meaning “the game of the palm” because it was originally played with the hand, and what was called Real Tennis in Britain, Court Tennis in the United States and Royal Tennis in Australia. The version of Royal tennis as played today can be traced back at least as far as 1490 in Urrungne France, which is possibly the oldest court in the world (Garnett, 1999). The modern version of tennis today (official name of Lawn Tennis) can be traced back to 1858, with the first lawn tennis club formed in Leamington in 1872, and the first official championships played in 1877 at Wimbledon. Tennis was a founding sport in the first modern Olympics in 1896 but was withdrawn in 1928 over disputes concerning the definition of an amateur. Returning as a demonstration sport in Los Angeles in 1984, it was reinstated as a full medal sport in Seoul in 1988. In 2010, there were 65 tournaments on the main tour in men's singles comprising of 4 Grand Slam, 10 Masters and 51 World Tour events. In 2010, there were 57 tournaments on the main tour in women's singles comprising of 4 Grand Slam, 11 Premier and 42 International events. Note that the Olympics is classified as a World Tour event in men's singles and an International event in women's singles.

The origins of the scoring system are more difficult to pin down that the antecedents of the game itself. Various references to the 0, 15, 30, 45 game structure have been noted during the 15th century, including a poem about the Battle of Agincout written in 1415. Originally it seems that the scoring was in fifteens going 15, 30, 45 but over time, instead of saying "forty-five", people started to say forty" for short and eventually was adopted as the official terminology. No-one really knows why counting in 15's originated. One hypothesis is that it is based around the clock face, the target score of 60 being a complete revolution of the minute hand. Another hypothesis which appears to be more plausible points to a French origin, since in the early middle ages, 60 was a key number in France in the same way that 100 is today. Most sports including tennis were played for money in the Middle Ages. There were laws in nearby Germany in the 14th and late 13th centuries that forbade stakes greater than sixty "deniers" and it happens that at about the same time there was a coin in circulation called a "gros denier tournois" which was worth 15 deniers. Perhaps the French tennis playing public was playing for one "gros denier tournois" per point up to the maximum stake of sixty deniers for a game (Gillmeister, 1997). The idea of "deuce" was introduced (at least as far as 1490) with a simple explanation - to ensure that the game could not be won by a one-point difference in players' scores. Hence deuce was derived from the French "a deux du jeu" - two points away from game.

2.1 Royal Tennis

The fixtures for professional Royal Tennis events can be found at the International Real Tennis Professional Association, and include each year the Australian Open, US Open, French Open and British Open, as well as the World Championships (every two years). The structure of a game which was recognized in 15th century tennis is still being used in Royal Tennis today and has carried through to Lawn Tennis. A player needs four points to win the game. If the score reaches deuce the game continues indefinitely until a player is two points ahead and wins the game. Sets however are played first-to-six games win the set, even if the opponent has five games. A match is typically best-of-three sets, except for the major open tournaments, in which matches are best-of-five sets. One fault is allowed on serve and a double fault results in the server losing the point. The spin of a racket (or toss of a coin) is used to decide the server of that match. As documented in the Laws of Real Tennis Australia (Garnett, 1999) –

during a match the players shall change sides when two Chases have been made or when any player is at forty or advantage and one Chase has been made.

2.2 Lawn Tennis – singles

The major change in the scoring structure from Royal Tennis to Lawn Tennis is in the requirements to win a set. Originally in Lawn Tennis all sets were advantage where a player needed six games to win the set. If the score reached five games all, the set continued indefinitely until a player was two games ahead and won the set. Another change from Royal Tennis is the rotation of service, where players rotate service after the completion of each game. One fault is allowed on serve and a double fault results in the server losing the point. The spin of a racket (or toss of a coin) is used to decide the server of that match. Men play best-of-five sets in grand slam matches and the Olympic Games gold medal match, and best-of-three sets in other matches on the main tour. Women always play best-of-3 sets.

This scoring system survived unchanged throughout the amateur era until 1968 when tennis was opened up to professional players, and tournaments became major television events. The tiebreak was invented by Jimmy Van Alen in 1965 to reduce the length of matches, and in 1970, the US Open became the first of the Grand Slam tournaments to use the tiebreak set, by originally playing a nine-point shootout (sudden death at 4-4) at 8 games all in the set (with the exception of the final set which was advantage). When the tiebreak was introduced at Wimbledon in 1971, it was called the 8-8 tiebreak, i.e., the first player to reach nine points with a lead of two. In 1979, the tiebreak was introduced when a set score reached 6 games all, as is the case today. The serving in a tiebreak is a rotation process where one player serves the first point, and serving then alternates after every two points. Among the Grand Slams, only the US Open today uses the tiebreak in the final set; the Australian Open, French Open and Wimbledon instead play an advantage final set. Olympic Games matches also use an advantage final set in both men's and women's singles matches.

2.3 Lawn Tennis – doubles

A best-of-three set match structure, where all sets are tiebreak sets are currently used in men's doubles events for the Australian Open, French Open and the US Open. A best-of-five set match structure is used in men's doubles events for Wimbledon where the deciding fifth set is advantage. A best-of-three set match structure is used in all grand slam women's doubles events where all tiebreak sets are used for the Australian Open, French Open and US Open, and an advantage deciding third set is used for Wimbledon. Olympic Games men's and women's doubles matches also use a best-of-three set match, where the deciding set is advantage.

In 2001 the Australian Open replaced the final set of mixed doubles with a match tie-break (first to 10 points and win by 2 points wins the match). Despite some criticism of the change by fans and former pros, the US Open and the French Open have since gone on to join the Australian Open in using the same format for mixed doubles. Wimbledon continues to play a traditional best-of-three set match with the final set being an advantage set.

In 2006 there was a change to the best-of-three sets scoring system used for doubles in men's and women's doubles tournaments on the main tour (excluding grand slam events). The main objectives of the change would appear to have been to reduce somewhat the average length of a match, to play matches that have a more predictable duration, and to reduce the likelihood of 'long' matches. The previous system was a best-of-three sets match structure, where all sets were tiebreak sets and games where standard "deuce" games. The system now used for these tournaments is a best-of-three sets system with the first two tiebreak sets using no-ad games, and the third set being simply a first-to-ten point's match-deciding tiebreak game. In no-ad games the next player to win the point if deuce is obtained is the winner.

3. Factors causing long matches

Long matches can cause problems for tournaments, can be unfair in the tournament setting and can also lead to injuries to the participating players (Barnett et al. 2006, Barnett and Pollard 2006). One factor that can lead to long matches is the use of the advantage set as the fifth set, as in the Australian Open, the French Open and Wimbledon. An advantage set requires a player to break serve to win the set, and this can be difficult to obtain when both players are serving a high percentage of first serves in with a "large" proportion resulting in aces. Other factors are long rallies and a greater than average number of points per game. These tend to occur more frequently on the slower court surfaces such as at the French Open.

3.1 Percentage of points won on serve

Table 1 represents the set score distribution for matches played in men's grand slam events across different time periods. The proportion of tiebreak games (7-6 score lines) increased from 11.4% in 1978-1982, to 12.1% in 1995-1999 and to 14.0% in 2000-2004. It is well documented (Barnett and Clarke, 2005) that the set score distribution is dependent on the percentage of points won on serve, such that increasing the percentage of points won on serve for both players will increase the proportion of tiebreaks played. This provides some justification that the percentage of points won on serve has increased in the last 30 years. A likely reason for this increase in serving performance is in the tennis racket. Haake et al (2007) concludes that if a player used rackets and balls from the 1870's, 1970's and 2007, then serve speeds would have increased by 17.5% since the 1870's, with a quarter of the change coming since the 1970's.

Score	1978-1982	1995-1999	2000-2004
6-0	6.3%	3.6%	3.9%
6-1	12.6%	10.9%	10.6%
6-2	20.7%	19.4%	18.1%
6-3	20.7%	21.9%	22.3%
6-4	19.8%	24.0%	22.7%
7-5	8.4%	8.1%	8.3%
7-6	11.4%	12.1%	14.0%

Table 1: Set score distribution for matches played in men's grand slam events across different time periods

Table 2 represents the percentage of points won on serve for men's and women's singles and doubles events in the 2001 Wimbledon, Australian Open and French Open. Overall, the results indicate that men win a higher percentage of points on serve compared to women for singles and doubles, and at all three grand slams. The results also indicate that for men and women, a higher percentage of points on serve occur at doubles events compared to singles events (at the same venue) and that there is an ordering of points won on serve with Wimbledon (grass) the highest followed by the Australian Open (hard) followed by the French Open (clay). This agrees with the results of Barnett and Pollard (2007).

Tables 3 and 4 represent the proportion of matches where the percentages of points won on serve for both players in men's (table 3) and women's (table 4) grand slam singles events in 2010 are greater than a specified amount. For example 7.2% of men's singles matches played at Wimbledon 2010 had both players winning more than 70% of points on serve. It can be shown that with a 70% chance of winning a point on serve, the chance of winning a game on serve is 90.1% (Barnett and Clarke, 2005). With the advantage final set used at Wimbledon, Australian Open and the French Open, this gives some justification to show why long matches can occur in men's singles grand slam events.

Event	Wimbledon 2001	Australian Open 2001	French Open 2001
Men’s singles	64.5%	61.9%	60.1%
Men’s doubles		62.9%	
Women’s singles	57.1%	54.9%	54.1%
Women’s doubles		55.4%	
Mixed doubles		63.0%	

Table 2: Percentage of points won on serve for men’s and women’s singles and doubles events in the 2001 Wimbledon, Australian Open and French Open

Serving Percentages	Wimbledon 2010	US Open 2010	Australian Open 2010	French Open 2010
>70%	7.2%	3.1%	2.4%	0.8%
>65%	29.6%	7.9%	9.6%	7.1%
>60%	66.4%	39.4%	36.8%	30.7%
>55%	91.2%	71.7%	72.0%	55.1%
>50%	98.4%	89.0%	86.4%	81.1%

Table 3: Proportion of matches represented by serving percentages at men’s grand slam singles events in 2010

Serving Percentages	Wimbledon 2010	US Open 2010	Australian Open 2010	French Open 2010
>70%	0.8%	0.0%	0.0%	0.0%
>65%	3.2%	0.8%	0.8%	0.8%
>60%	19.0%	8.7%	7.1%	1.6%
>55%	48.4%	21.4%	31.5%	15.0%
>50%	79.4%	40.5%	63.8%	43.3%

Table 4: Proportion of matches represented by serving percentage at women’s grand slam singles events in 2010

In recent years there have been a number of grand slam matches decided in long fifth sets. In the third round of the 2000 Wimbledon men's singles, Philippoussis defeated Schalken 20–18 in the fifth set. Ivanisevic defeated Krajicek 15–13 in the semi-finals of Wimbledon in 1998. In the quarter-finals of the 2003 Australian Open men's singles, Andy Roddick defeated Younes El Aynaoui 21–19 in the fifth set, a match taking 83 games to complete and lasting a total duration of 5 h. The night session containing this long match required the following match to start at 1 am. Barnett and Clarke (2005) give a detailed analysis of this Roddick versus El Aynaoui match and conclude that whenever two players with dominant serves but relatively poor returns of serve meet, there is always a chance that if the match reaches a fifth set, it can go on for a long period of time. They showed from pre-match predictions that this match was likely to go longer than any other men's singles match played at the 2003 Australian Open.

The longest professional tennis match, in terms of both time and total games, was the more recent Wimbledon 2010 first-round match between Nicholas Mahut and John Isner. It lasted 183 games and required 11 hours and 5 minutes of playing time. Even with the introduction of a tiebreak set in 1970 long matches still occur and records of long matches can still be broken. Table 5 represents the fifth set match statistics, where the 1st Serve % of 67% and 74% for Mahut and Isner respectively, are both higher than the ATP tour average on grass of 61.9% (Bedford et al, 2010). It can be observed from the Aces that 22.4% and 23.1% of points served by Mahut and Isner respectively resulted in an ace. In the fifth set, Mahut won 82.5% of points on serve and Isner won 79.6% of points on serve, which are both significantly higher than the ATP tour average on grass of 65.5% (Bedford et al, 2010). It is also worth noting the relatively small number of break point opportunities for both players – 2 for Mahut and 5 for Isner.

Set 5 summary	Mahut	Isner
1st Serve %	231 of 343 = 67%	272 of 368=74%
Aces	77	85
Double Faults	12	4
Unforced Errors	25	33
Winning % on 1st Serve	208 of 231=90%	223 of 272=82%
Winning % on 2nd Serve	73 of 112=65%	61 of 96=64%
Winners	183	182
Receiving Points Won	84 of 372=23%	62 of 355=17%
Break Point Conversations	0 of 2=0%	1 of 5=20%
Net Approaches	82 of 106=77%	74 of 105=70%
Total Points Won	365	346

Table 5: Fifth set match statistics for Wimbledon 2010 first-round match between Nicholas Mahut and John Isner

3.2 Length of time to play a point

The longest match by time played at the Australian Open was between Rafael Nadal and Fernando Verdasco in a 2009 semifinal match, lasting for 5 hours and 14 minutes. The advantage fifth set score line was only 6-4, which suggests that best-of-five set matches with all tiebreak sets (such as the system used in the US Open) can still produce long matches. There was a match between Arnaud Clement and Fabrice Santoro played at the 2004 French Open that lasted for 6 hours 36 minutes, and is the longest match by time played at the French Open and the longest match by the number of games played at the French Open since the introduction of the tiebreak set. Although only 71 games were played in this match, the time duration was longer than the Roddick versus El Aynaoui match played at the 2003 Australian Open. Table 6 represents the percentage of points won on serve for each player for each set and the time taken to complete each set with the corresponding game score. It took an average time of 55.75 minutes to play each set in the first four tiebreak sets. These relatively long tiebreak sets must be due to the length of time to play each game, which is a combination of the number of points played in the game and the length of time to play each point. The average percentage of points won on serve for each player in the first four sets is 54.5% for Clement and 56.0% for Santoro, which are both less than the ATP tour average of 61.6%. Since there is a lack of dominance on serve, it is most likely that the length of time to play each point is higher than the ATP tour average time to play each point. Notice that the percentage of points won on serves for each player in the fifth advantage set is 64%, which is at least as high as any of the other sets, contributing to the 30 games and 173 minutes to play the final set.

Table 7 was taken from Croucher (1998) and gives the statistics on the time and number of points for three Grand Slam championships in 1991. For the men, the average number of points played is significantly higher at the French Open on clay. This is a result of the percentage of points won on serve being less on clay than the surfaces of grass and hard, leading to longer games. For both men and women, the average time to play a point and the average play time per hour is significantly higher at the French Open on clay.

	Serving statistics (%)			
	Clement	Santoro	Time (min)	Score
Set 1	56	61	51	4-6
Set 2	50	64	46	3-6
Set 3	55	56	74	7-6
Set 4	57	43	52	6-3
Set 5	64	64	173	14-16
Match	58	60	396	

Table 6: Statistics for each set obtained from the Clement versus Santoro match played at the 2004 French Open

	French Open (clay)	US Open (hard)	Wimbledon (grass)
Men			
Avg. number of points	279	155	214
Avg. time per point	10.0s	7.6s	2.6s
Avg. play time per hour	14min 56s	9min 58s	3min 42s
Women			
Avg. number of points	122	119	200
Avg. time per point	11.0s	5.3s	5.9s
Avg. play time per hour	15min 43s	9min 41s	9min 18s

Table 7: Time and number of points for three Grand Slam championships in 1991

Based on the results from Table 7 and the likelihood of long matches in terms of time on clay and hard court, there is evidence to suggest that a tiebreak set may be too long in a best-of-five set match. One method to reduce the length of a tiebreak set is to use no-ad games as currently used in a range of doubles events. The characteristics of such scoring systems will be investigated in section 4.

4. Scoring Systems

4.1 Characteristics

Many papers appear in the literature in comparing scoring systems for tennis (Brown et al 2008a, Pollard et al 2007, Pollard and Noble 2004). Typical characteristics to make comparisons for these scoring systems are:

(i) P(A wins) where A is the better player/pair

(ii) Mean number of points in the match

(iii) Standard deviation of the number of points in the match

(iv) Efficiency of the scoring system

(v) 98% point of the cumulative distribution of the number of points

Table 8 lists the scoring systems to be analyzed by representing the game, set and match structure. The 50-40 game as outlined in Pollard and Noble (2004) is such that to win the game, the server requires four points and the receiver requires three points. There is at most six points played in this type of game.

System	Event	Game	Set	Match
1	Royal Tennis	Deuce	First-to-six	3 or 5 sets
2	Pre-1970 men's Lawn Tennis	Deuce	All advantage	5 sets
3	Pre-1970 women's Lawn Tennis	Deuce	All advantage	3 sets
4	Aust./French/Wimb. men's singles Olympics men's singles (Gold medal) Olympics men's doubles (Gold medal) Wimb. men's doubles	Deuce	Final set advantage	5 sets
5	Aust./French/Wimb. women's singles Olympics men's and women's singles Olympics men's and women's doubles Wimb. women's doubles Wimb. mixed doubles	Deuce	Final set advantage	3 sets
6	US Open men's singles Aust./French/US Open men's doubles	Deuce	All tiebreak	5 sets
7	US Open women's singles Aust./French/US Open women's doubles Men's and women's singles Pre-2006 men's and women's doubles	Deuce	All tiebreak	3 sets
8	Aust/French/US Open mixed doubles	Deuce	Final set super tiebreak game	3 sets

9	Men's and women's doubles	No-ad	Final set super tiebreak game	3 sets
10	Alternate men's grand slam singles	No-ad	All tiebreak	5 sets
11	Alternate men's grand slam singles	50-40	All tiebreak	5 sets
12	Alternate men's grand slam singles	50-40	Final set advantage	5 sets
13	Alternate non grand slam doubles	50-40	Final set super tiebreak game	3 sets

Table 8: Types of scoring systems for men's and women's singles and doubles matches on the main tour

4.2 Results

Tables 9, 10 and 11 give numerical results for the five characteristics (i)-(v) for scoring systems 2-13 given in table 8. The statistical measures of p_a and p_b (the probabilities of player/pair A and player/pair B winning a point on their respective serve) are given to represent matches at the 'strong-serving' end (p_a = 0.77, p_b = 0.73), and at the 'weaker-serving' end (p_a = 0.62, p_b=0.58) (Brown et al 2008a). For the 12 scoring systems to be analyzed, the probability that the better player/pair wins the match, and the mean and higher moments of the number of points played, were evaluated using recursive methods (Brown et al., 2008b). The process used to estimate the distribution of the number of points in a match was based on using the Normal Power approximation (Pollard et al., 2007) and simulated results.

	2		3		4		5	
	p_a=0.77 p_b=0.73	p_a=0.62 p_b=0.58	p_a=0.77 p_b=0.73	p_a=0.62 p_b=0.58	p_a=0.77 p_b=0.73	p_a=0.62 p_b=0.58	p_a=0.77 p_b=0.73	p_a=0.62 p_b=0.58
(i)	0.778	0.752	0.730	0.707	0.723	0.743	0.690	0.701
(ii)	484.3	274.2	298.2	168.3	290.3	262.1	192.1	161.6
(iii)	218.2	73.8	165.0	51.6	99.5	63.6	93.1	44.6
(iv)	0.34	0.61	0.36	0.65	0.34	0.59	0.37	0.63
(v)	1049	442	750	295	582	395	480	261

Table 9: The five statistics for scoring systems 2-5

	6		7		8		9	
	p_a=0.77 p_b=0.73	p_a=0.62 p_b=0.58	p_a=0.77 p_b=0.73	p_a=0.62 p_b=0.58	p_a=0.77 p_b=0.73	p_a=0.62 p_b=0.58	p_a=0.77 p_b=0.73	p_a=0.62 p_b=0.58
(i)	0.708	0.741	0.669	0.697	0.656	0.670	0.658	0.658
(ii)	272.0	261.0	166.3	160.0	142.8	137.8	131.5	122.0
(iii)	60.7	61.6	40.3	41.4	21.8	24.5	20.5	20.5
(iv)	0.32	0.58	0.34	0.62	0.33	0.52	0.37	0.51
(v)	385	383	243	246	187	191	174	166

Table 10: The five statistics for scoring systems 6-9

	10		11		12		13	
	p_a=0.77 p_b=0.73	p_a=0.62 p_b=0.58	p_a=0.77 p_b=0.73	p_a=0.62 p_b=0.58	p_a=0.77 p_b=0.73	p_a=0.62 p_b=0.58	p_a=0.77 p_b=0.73	p_a=0.62 p_b=0.58
(i)	0.712	0.721	0.727	0.715	0.730	0.715	0.667	0.655
(ii)	248.9	229.9	210.0	196.3	211.7	196.6	112.5	105.2
(iii)	55.7	52.7	48.6	45.7	52.0	46.3	19.9	18.8
(iv)	0.36	0.55	0.50	0.60	0.51	0.61	0.48	0.56
(v)	354	333	306	286	323	290	158	146

Table 11: The five statistics for scoring systems 10-13

4.3 Discussion

Royal Tennis, at least as far back as 1490, used a scoring structure such that matches were the best-of-five or best-of-three sets, sets were first-to-six and games were deuce games. It would appear that when Lawn Tennis was introduced to Wimbledon in 1877 that the service could be an advantage (discussed in the introduction), and therefore using first-to-six games could present a significant advantage to the player serving first in the match. In a similar way that two consecutive points are required to win a deuce game if scores are level at three points all, a set in Lawn Tennis would require a player to win two consecutive games if scores are level at five games all.

The introduction of the tiebreak game in 1970 was to reduce the length of matches and advantage sets are no longer played under the rules of the United States Tennis Association.

However, advantage sets are still used in the final set in the Australian Open, French Open, Wimbledon and the Olympics. The US Open is the only Grand Slam to use a tiebreak in the final set. There was evidence to show that serving performance has increased in the last 30 years due to changes in racket technology (Haake et al. 2007). This has resulted in long matches in recent times from an advantage final set. More recently, the longest match in history was recorded at 11 hours and 5 minutes. From table 3, there was evidence to show that long final advantage sets are a reasonable possibility in men's singles matches. From table 9 it can be observed that the 98% point of the cumulative distribution of the number of points for system 4 when p_a=0.77 and p_b=0.73 is 582 and when p_a=0.62 and p_b=0.58 is 395. In comparison from table 10, the 98% point of the cumulative distribution of the number of points for system 6 when p_a=0.77 and p_b=0.73 is 385 and when p_a=0.62 and p_b=0.58 is 383. Given that two 'strong' servers have led to long matches in men's tennis, it could be argued that system 6 is preferable to system 4 in men's singles and doubles matches.

From section 3, it was observed that a tiebreak set may be too long in a best-of-five set match, as a result of the amount of time to play a point (particularly on clay) and the 'lack' of dominance on serve leading to a relative 'large' number of points played in a game. Long matches played at the Australian and French Open was used to support this observation. One method to reduce the length of a tiebreak set is to use no-ad games as currently used in a range of doubles events. By comparing the characteristics of system 6 to system 10; the mean, standard deviation and the 98% point of the cumulative distribution of the number of points are reduced for system 10 for when p_a=0.77 and p_b=0.73, and for when p_a=0.62 and p_b=0.58. System 10 is also more efficient that system 6. The 50-40 game as outlined in Pollard and Noble (2004) is such that to win the game, the server requires four points and the receiver requires three points. When comparing system 10 to system 11; the mean, standard deviation and the 98% point of the cumulative distribution of the number of points are reduced for system 11 for when p_a=0.77 and p_b=0.73, and for when p_a=0.62 and p_b=0.58. System 11 is also more efficient that system 10. Based on the above, it could be argued that systems 11 and 10 are preferable to system 6 and system 11 is preferable to system 10. The characteristics are very similar when comparing system 11 to system 12, and demonstrates that an advantage final set is 'reasonable' in reducing the likelihood of long matches occurring by changing the game structure of a standard deuce game to a 50-40 game.

From section 2, in 2006 there was a change to the best-of-three sets scoring system used for doubles in a range of professional men's and women's tournaments (excludes grand slam events). The main objectives of the change would appear to have been to reduce somewhat the average length of a match, to play matches that have a more predictable duration, and to reduce the likelihood of 'long' matches. Pollard et al. (2007) compare five alternative scoring systems including the 50-40 game and conclude that on statistical grounds these systems would appear to be legitimate alternatives to the current system. By comparing the characteristics between system 9 and 13, it could be argued that system 13 is preferable to system 9.

It was shown in table 2 that the average percentage of points won on serve in mixed doubles at the 2011 Australian Open was 63.0%, which is higher than both men's singles (61.9%) and men's doubles (62.9%). For this reason, it could be argued that system 8 is preferable to system 5.

Based on the results and discussion, the following recommendations are obtained for men's and women's singles and doubles events and represented in table 12.

	Match	Set	Game
Men's singles (grand slam/Olympics)	5 sets	Final set advantage	50-40
Men's singles	3 sets	All tiebreak	Deuce
Men's doubles (grand slam/Olympics)	3 sets	All tiebreak	Deuce
Men's doubles	3 sets	Final set super tiebreak game	50-40
Women's singles (grand slam/Olympics)	3 sets	Final set advantage	Deuce
Women's singles	3 sets	All tiebreak	Deuce
Women's doubles (grand slam/Olympics)	3 sets	All tiebreak	Deuce
Women's doubles	3 sets	Final set super tiebreak game	50-40
Mixed doubles (grand slam)	3 sets	Final set super tiebreak game	Deuce

Table 12: Recommendations of scoring systems for men's and women's singles and doubles events

5. Conclusions

The history of tennis scoring systems, possible factors causing long matches, examples of long matches that have occurred in recent times and numerical results for characteristics of scoring systems are given in this paper. Based on this information, recommendations of scoring systems are given for men's and women's singles and doubles events that could be used for today. These recommendations show that men's and women's doubles matches (excluding grand slam events) could replace the current no-ad game with a 50-40 game and that men's singles grand slam and Olympic games matches could use a best-of-5 set match structure with a final set advantage by replacing the current deuce game with a 50-40 game.

References

Barnett T and Clarke SR (2005). Combining player statistics to predict outcomes of tennis matches. IMA Journal of Management Mathematics. 16 (2), 113-120.

Barnett T, Brown A and Pollard G (2006). Reducing the likelihood of long tennis matches. Journal of Sports Science & Medicine. 5(4), 567-574.

Barnett T and Pollard G (2006). Reducing injuries by substantially decreasing the likelihood of long tennis matches. Medicine and Science in Tennis. 11 (2), 10-11.

Barnett T and Pollard G (2007). How the tennis court surface affects player performance and injuries. Medicine and Science in Tennis. 12 (1), 34-37.

Bedford A, Barnett T, Pollard GH and Pollard GN (2010). How the interpretation of match statistics affects player performance. Journal of Medicine and Science in Tennis 15(2), 23-27.

Brown A, Barnett T, Pollard GH, Lisle I and Pollard GN (2008a). The characteristics of various men's tennis doubles scoring systems. In proceedings of the Ninth Australian Conference on Mathematics and Computers in Sport.

Brown A, Barnett T and Pollard G (2008b). A recursion method for evaluating the moments of a nested scoring system. In proceedings of the Ninth Australian Conference on Mathematics and Computers in Sport.

Croucher JS (1998). Developing strategies in tennis, In Statistics in Sport, J. Bennett ed. London: Arnold, 157–170.

Garnett M (1999). A chase down under: a history of royal tennis in Australia. Historical Publications.

Gillmeister H (1997). Tennis: a cultural history. Leicester University Press.

Haake S, Allen T, Choppin S and Goodwill S (2007). The evolution of the tennis racket and its effect on serve speed. Tennis Science and Technology 3 . Roehampton. S. Miller and J. Capel-Davies (eds), 257-271.

Pollard GH, Barnett T, Brown A and Pollard GN (2007). Some alternative men's doubles scoring systems. Tennis Science and Technology 3 . Roehampton. S. Miller and J. Capel-Davies (eds), 301-309.

Pollard G and Noble K (2004). The benefits of a new scoring system in tennis: the 50-40 game. In R. Morton and S. Ganesalingam (eds.) Proceeding of the Seventh Australasian Conference on Mathematics and Computers in Sport. Massey University, Massey, New Zealand. 262-265.

Challenge System

Introduction

The new challenge system for close line calls in tennis has been used on the ATP and WTA tour for grand slam events since the 2006 US Open, and was designed to increase fairness for players by obtaining accurate line calls and enhance spectator interest through video technology. In the current system, players have unlimited opportunity to challenge, but once three incorrect challenges are made in a set, they cannot challenge again until the next set. If the set goes to a tiebreak game, players are given additional opportunities to challenge (usually one extra). If the match is tied at six games all in an advantage set, the counter is reset with both players again having a limit of up to three incorrect challenges in the next 12 games, and this resetting process is repeated after every 12 games.

Strategies as to when players should challenge have recently appeared in the literature. Pollard et al. (2010) show that challenge decisions are based on the rate at which challenges occur, the expected number of points remaining in the set, the number of challenges remaining in the set, the probability of the challenge decision being successful and the importance of the point to winning the set. Clarke and Norman (2010) apply dynamic programming to the challenge system to investigate the optimal challenge strategy and obtain some general rules.

There appears to be problems with the current challenge system:

- Firstly, both of the above articles show that early in the set a player needs to decide whether to challenge, or save challenges to later on in the set when the points are typically more important. Having to make such decisions is completely irrelevant to the game of tennis itself. The aim of the contest is to find the better player, and not to favour the player who is luckier within, or better at playing the challenge system. This is reflected by an article *Replay System Becomes Part of Players' Strategies* in The New York Times by Greg Bishop during the 2009 US Open. http://www.nytimes.com/2009/09/11/sports/tennis/11challenges.html

- Secondly, a player can run out of challenges because that particular set has a lot of balls that go close to the lines. This is perhaps particularly true in men's singles and men's doubles. The problem is exacerbated when each player does not have a similar number of challenges. A player who plays more balls near the lines is disadvantaged relatively. The player who, by chance has the need for more challenges, is disadvantaged.

- Thirdly, it would appear to be disappointing for the player and the spectators when that player runs out of challenges, the point is very important, and a challenge would have a clear likelihood of success. What is the chance that a grand slam final will be 'messed up' by an umpire making a wrong call and the player having run out of challenges, and subsequently losing the final when he might well have won it otherwise? This would be a very bad result for the player, the umpire and the game. Maybe this probability is not quite as small as some people might expect.

Importance of points

Morris (1977) defines the importance of a point for winning a game (I_{PG}) as the probability that the server wins the game given he wins the next point minus the probability that the server wins the game given he loses the next point. Table 1 gives the importance of points to winning the game when the server has a 0.62 probability of winning a point on serve, and shows that 30-40 and Ad-Out are the most important points in the game. In a similar way, we can define the importance of a game to winning a set and the importance of a set to winning a match. Table 2 gives the importance of games to winning a tiebreak set (I_{GS}) for player A serving. Player A and Player B were assigned point probabilities of 0.62 and 0.60 respectively to reflect overall averages in men's tennis. It is clear that every point is equally important for both players. Table 2 shows that the tiebreak game has the highest importance of 1.00, as the winner of this game wins the set. Similarly, table 3 gives the importance of sets to winning a best-of-5 set match (I_{SM}) and shows that the deciding set at 2 sets-all has the highest importance of 1.00, as the winner of this set win the match. Morris (1977) derived the following useful multiplicative result to obtain the importance of a point to winning the match (I_{PM}): For any point of any game of any set, $I_{PM} = I_{PG} * I_{GS} * I_{SM}$.

The definition of importance of a point in a match is a way of stating how much difference will result in the outcome of the match depending on whether a point is won or lost. In the context of a challenge system, importance of a point in a match can be viewed by how much percentage error will occur if a wrong decision is made. For example, suppose the score in a best-of-5 set match (all tiebreak sets) is 2-2 in sets, 5-5 in games and 30-30 in points and player A is currently serving. Suppose player A is winning 62% on serve and player B is winning 60% on serve. Using a Markov Chain model (Barnett and Clarke, 2005), player A has a 51.5% chance of winning the match from that position. If player A won the point, then his chance of winning the match would be 60.3%: whereas if player A lost the point then his chance of winning the match would be 37.3%. Therefore the importance of the point in the match is given as 60.3%-37.3%=23.0%. If a wrong decision was made at that particular point in the match, then it would cost one of the players 23 percentage points in their chance of winning the match.

	Receiver's score					
		0	15	30	40	Ad
Server's score	0	0.25	0.34	0.38	0.28	
	15	0.19	0.31	0.45	0.45	
	30	0.11	0.23	0.45	0.73	
	40	0.04	0.10	0.27	0.45	0.73
	Ad				0.27	

Table 1: Importance of points to winning a game when the server has a 0.62 probability of winning a point on serve

	Player B's score							
		0	1	2	3	4	5	6
Player A's score	0	0.29	0.29	0.22	0.18	0.06	0.02	
	1	0.26	0.32	0.33	0.21	0.16	0.03	
	2	0.25	0.29	0.36	0.37	0.20	0.11	
	3	0.13	0.27	0.33	0.42	0.43	0.14	
	4	0.08	0.11	0.30	0.38	0.52	0.54	
	5	0.01	0.06	0.08	0.34	0.46	0.52	0.53
	6						0.47	1.00

Table 2: Importance of games to winning a tiebreak set when player A and player B have a 0.62 and 0.60 probability of winning a point on service respectively and player A is serving

	B's score			
		0	1	2
A's score	0	0.36	0.42	0.32
	1	0.32	0.49	0.57
	2	0.18	0.43	1.00

Table 3: Importance of sets to winning a best-of-5 set match when player A and player B have a 0.62 and 0.60 probability of winning a point on service respectively

Proposed new challenge system

It is proposed that the present challenge rule is modified in one way. Namely, that a player is allowed to challenge on points with sufficiently large importance, without risking that player's challenge point total.

Suppose the threshold value on when a player can always challenge a line call was given by the importance of the point in the match at 2 sets-all, 3 games-all and player A serving. This is calculated as 0.247*0.420*1.00=0.104 when player A and player B have a 0.62 and 0.60 probability of winning a point on serve respectively. Then a player can always challenge at 2 sets-all and 3 games-all, only if the point score in the game has an importance of at least 0.247. This occurs at 30-40 or Ad-Out (I_{PG}=0.727), 15-40 (I_{PG}=0.451), 15-30 (I_{PG}=0.447), 30-30 or deuce (I_{PG}=0.446), 0-30 (I_{PG}=0.384), 0-15 (I_{PG}=0.341), 15-15 (I_{PG}=0.315), 0-40 (I_{PG}=0.279), 40-30 or Ad-In (I_{PG}=0.273) and 0-0 (I_{PG}=0.247). This is represented in table 4 for a range of game scores in the deciding set, where an X indicates that a challenge is always allowable by both players. Note that a player can challenge at 2 sets-all and 6 games-all (tiebreak game), only if the point score has an importance of at least 0.104. This occurs for the majority of points in the tiebreak game, as expected.

	Score line at 2 sets-all (player A serving)					
Point score	0-0	1-1	2-2	3-3	4-4	5-5
30-40 or Ad-Out	X	X	X	X	X	X
15-40	X	X	X	X	X	X
15-30	X	X	X	X	X	X
30-30 or Deuce	X	X	X	X	X	X
0-30	X	X	X	X	X	X
0-15		X	X	X	X	X
15-15			X	X	X	X
0-40				X	X	X
40-30 or Ad-In				X	X	X
0-0				X	X	X
30-15					X	X
15-0, 30-0, 40-15 or 40-0						

Table 4: Indication as to whether a player can always challenge on a particular point in a match for a range of game scores in the deciding set given that the threshold value is given as 0.104

References

Barnett T and Clarke SR (2005), Combining player statistics to predict outcomes of tennis matches. IMA Journal of Management Mathematics. 16 (2), 113-120.

Clarke S and Norman J (2010), An introductory analysis of challenges in tennis. Proceedings of the Tenth Australasian Conference on Mathematics and Computers in Sport. Edited by A Bedford and M Ovens. 43-48.

Morris C (1977), The most important points in tennis. In Optimal Strategies in Sports, Volume 5 in Studies in Management Science and Systems, Edited by S.P. Ladany and R.E. Machol, Amsterdam: North-Holland Publishing Company, 131-140.

Pollard GN, Pollard GH, Barnett T and Zeleznikow J (2010). Applying strategies to the tennis challenge system. Journal of Medicine and Science in Tennis 15(1), 12-15.

Serving Strategies

Introduction

Analyzing risk-taking strategies in tennis is complicated. There has been a tendency to analyze risk-taking on the serve more often than other shots. This seems reasonable as the serve is the first shot to be played and therefore simplifies the analysis by not having to consider previous shots in the rally. Barnett et al. (2008) analyzed the situation where players may choose to serve two fast serves by taking into account the type of court surface, and the serving and receiving capabilities of both players. Pollard et al. (2009) extend on this model by allowing for the possibility of players changing serving strategies throughout the match in progress. Consideration of the ideal that a continuum amount of risk is available to players on their serve has further revealed a higher risk first serve and a lower risk second serve strategy as being optimal in most practical situations (Pollard et al., 2007). Pollard (2008) also analyzed the situation in which a medium risk serve (somewhere between a players 'typical' high risk first serve and low risk second serve) has a quadratic rather than linear outcome; one which gives greater weighting to the outcome of serving a high risk serve rather than the outcome of a low risk serve.

All of the above articles analyze the situation where the server is the only decision maker and therefore the optimal strategy will be a single strategy with certainty e.g. a player should always serve a 'typical' high risk first serve on both the first and second serves. When analyzing risk taking on serve by also taking into account whether the receiver is expecting a low or high risk second serve (known more generally as game theory), the optimal strategy can be a mixed strategy e.g. a player should serve a 'typical' high risk first serve 20% of the time on the second serve and a 'typical' low risk second serve 80% of the time on the second serve. This game theory scenario will be analyzed in this article and extended to include the 'importance' of points; where it is suggested for the server to take more risk on the more 'important' points i.e. 30-40 is shown to be the most 'important' point in a game.

Data Collection and Analysis

Match statistics from OnCourt (www.oncourt.info) can be obtained for the majority of ATP and WTA matches dating back to 2003. Using a customized program, the average serving and receiving statistics for each player on each surface were calculated, as well as the average head-to-head serving and receiving statistics between any two players.

Bedford et al. (2010) show how a range of statistics (such as the percentage of points won on serve and the 2nd Serve %) can be obtained from the broadcasted match statistics. Table 1 gives the match statistics broadcast from The Artois Championships in 2008 (played on grass) where Rafael Nadal defeated Andy Roddick in two straight sets. Notice that the Serving Points Won is not given directly in the table. This statistic can be derived from the Receiving Points Won such that Serving Points Won for Nadal and Roddick are 1-14/61=77.0% and 1-24/71

=66.2% respectively. Note that the Winning % on 1st Serve is conditional on the 1st Serve going in whereas the Winning % on the 2nd Serve is unconditional on the 2nd Serve going in. The Serving Points Won for Nadal and Roddick along with the Winning % on 1st Serve (uncond.), Winning % on 2nd Serve (cond.) and 2nd Serve % are given in table 2.

	Rafael Nadal	Andy Roddick
1st Serve %	45 of 61 = 73%	46 of 71 = 64%
Aces	7	14
Double Faults	0	3
Winning % on 1st Serve (cond.)	35 of 45 = 77%	34 of 46 = 73%
Winning % on 2nd Serve (uncond.)	12 of 16 = 75%	13 of 25 = 52%
Break Point Conversions	2 of 7 = 28%	0 of 4 = 0%
Receiving Points Won	24 of 71 = 33%	14 of 61 = 22%
Total Points Won	71	61

Table 1: Match statistics between Rafael Nadal and Andy Roddick at The Artois Championships in 2008

	Rafael Nadal	Andy Roddick
Serving Points Won	1-14/61=77.0%	1-24/71 =66.2%
Winning % on 1st Serve (uncond.)	(45/61)*(35/45)=57.4%	(46/71)*(34/46)=47.9%
Winning % on 2nd Serve (cond.)	(12/16)/(1-0/61)=75.0%	(13/25)/(1-3/71)=54.3%
2nd Serve %	1-0/61=100.0%	1-3/71=95.8%

Table 2: Calculated statistics between Rafael Nadal and Andy Roddick at The Artois Championships in 2008

Results

Scenario a)

The model developed in Barnett et al. (2008) is used to determine if the server can increase their chances of winning a point by serving high risk on the second serve. As outlined in the introduction this scenario is such that the server is the only decision maker and therefore the optimal strategy will be a single strategy with certainty.

The following definitions are given to obtain a high and low risk serve for each player:

- A high risk serve is a ‘typical’ first serve by a player and calculations are obtained by a player’s averaged percentage of points won on the first serve for a particular surface
- A low risk serve is a ‘typical’ second serve by a player and calculations are obtained by a player’s averaged percentage of points won on the second serve for a particular surface

Note the limitations in these definitions of a high and low risk serve in that to obtain a reasonable sample size a player's serving statistics is across all players (rather than just head-to-head against the opponent). Also a 'typical' first and second serve by each player may not be consistent across each match, but rather a player may be taking more 'risk' on the second serve on particular matches for example.

Let:
d_{hijs} = percentage of points won on high risk serves (unconditional) for player i, for when player i meets player j on surface s
d_{lijs} = percentage of points won on low risk serves (unconditional) for player i, for when player i meets player j on surface s

The following two serving strategies are defined:
Strategy 1 – high risk serve followed by a high risk serve
Strategy 2 – high risk serve followed by a low risk serve

Thus, player i should use Strategy 1 (two high risk serves) rather than Strategy 2 if $d_{hijs} > d_{lijs}$

An example of such a case is given in Barnett et al. (2008) between Andy Roddick (recognized as a 'strong' server) and Rafael Nadal (recognized as a 'strong' receiver), where the results from table 3 indicate that Roddick might be encouraged to serve high risk on both the first and second serve when playing Nadal on grass (since 0.535>0.512). However he should use a high risk first serve and low risk second serve when playing Nadal on both hard court (since 0.528<0.551) and clay (since 0.364<0.458). This example illustrates the fact that it can be important for players to identify the match statistics for themselves and their opponents – specific to court surfaces.

	Andy Roddick			Rafael Nadal		
Statistic	Grass	Hard	Clay	Grass	Hard	Clay
d_{lijs}	0.512	0.551	0.458	0.582	0.571	0.608
d_{hijs}	0.535	0.528	0.364	0.510	0.495	0.546
matches	37	99	17	24	72	72

Table 3: Serving and receiving statistics for Andy Roddick and Rafael Nadal

Scenario b)

The model developed in scenario a) is now extended by taking into account strategies on whether the receiver is expecting a low or high risk second serve. From table 3, where Roddick is serving against Nadal on hard court, Roddick is expected to win 55.1% of points on the second serve when serving low risk on the second serve and expected to win 52.8% of points on the second serve when serving high risk on the second serve. Suppose these percentages are based on whether Nadal on the return of serve is expecting a high or low risk second serve. For example, if Roddick was serving a low risk second serve and Nadal was expecting a low risk second serve, then the percentage won on the second serve for Roddick would likely be less than 55.1%. This is represented in table 4 below in a game theory matrix with the following observation. If Nadal was expecting a low risk second serve 50% of the time and a high risk second serve 50% of the time (indifferent between strategies), then Roddick should always serve a low risk second serve since ½*0.53 + ½*0.57=0.55 and ½*0.55 + ½*0.51=0.53. These results are in agreement with the earlier model from scenario a) where decisions of the opponent were not taken into account.

Using standard game theory techniques to solve this two-person zero-sum game; gives mixed strategies for Roddick of 50% low risk serve, 50% high risk serve and for Nadal of 75% expecting a low risk serve, 25% expecting a high risk serve. The outcome of the game with both players' adopting these mixed strategies is such that Roddick will win 54% of points on the second serve. If either player deviated from these strategies then the other player could capitalize by changing strategies accordingly. For example, if Roddick changed strategies to 80% low risk serve, 20% high risk serve, then Nadal could choose the strategy of 100% expecting low risk serve, for an outcome of Roddick to win 0.53*0.8+0.55*0.2=53.4% of points on the second serve.

		Nadal	
		expecting low risk serve	expecting high risk serve
Roddick	low risk serve	0.53	0.57
	high risk serve	0.55	0.51

Table 4: Game theory matrix of how much risk to take on the second serve in tennis

Scenario c)

The model developed in scenario a) is now extended to include the 'importance' of points. The results obtained also extend to the model developed in scenario b). Morris (1977) defines the 'importance' of a point for winning a game as the probability that the server wins the game given he wins the next point minus the probability that the server wins the game given he loses the next point. Table 5 gives the 'importance' of points to winning the game when the server has a 0.62 probability of winning a point on serve, and shows that 30-40 and Ad-Out are the most 'important' points in the game.

		Receiver's score				
		0	15	30	40	Ad
Server's score	0	0.25	0.34	0.38	0.28	
	15	0.19	0.31	0.45	0.45	
	30	0.11	0.23	0.45	0.73	
	40	0.04	0.10	0.27	0.45	0.73
	Ad				0.27	

Table 5: 'Importance' of points to winning a game when the server has a 0.62 probability of winning a point on serve

The following result follows from Klaassen and Magnus (2001), where it was established that a server's probability of winning a point decreases with the more 'important' points.

Player i should use Strategy 1 (two high risk serves) rather than Strategy 2 if $d_{hijs} > d^{\wedge}_{lijs}$. The superscript ^ is used as the server's probability of winning a point on a low risk serve is now conditional on the 'importance' of the point.

This is evidence to suggest that the server would be encouraged to take more risk on the more 'important' points.

Conclusion

The results obtained in this paper could be used by coaches to help determine how much risk their players should take on the second serve. By using the definitions of a high risk serve as a 'typical' first serve by each player and a low risk serve as a 'typical' second serve by each player, a model where the server was the only decision maker (does not take into account strategies on whether the receiver is expecting a low or high risk second serve) was formulated to determine how much risk a player should take on the second serve. An example was provided between Roddick and Nadal, where it was shown that Roddick might do slightly better when playing Nadal on grass by using two high risk serves rather than using a high risk first serve and a low risk second serve. By establishing a game theory model (by taking into account strategies on whether the receiver is expecting a low or high risk second serve) it was then shown that Roddick against Nadal on hard court could use mixed strategies on serving low and high risk on the second serve, even though the earlier model (that does not take into account strategies on whether the receiver is expecting a low or high risk second serve) indicates that Roddick should be serving low risk on every second serve with certainty for the entire match. Finally, consideration was given to the 'importance' of points which then pointed to the server being encouraged to take more risk on the more 'important' points.

References

Barnett T, Meyer D and Pollard G (2008). Applying match statistics to increase serving performance. *J Med Science Tennis,* 13(2):24-27.

Bedford A, Barnett T, Pollard GH and Pollard GN (2010). How the interpretation of match statistics affects player performance. Journal of Medicine and Science in Tennis 15(2), 23-27.

Klaassen and Magnus (2001). Are points in tennis independent and identically distributed? Evidence from a dynamic binary panel data model, Journal of the American Statistical Association 96, 500–509.

Morris C (1977). The most important points in tennis, In Optimal Strategies in Sports, S.P. Ladany and R.E. Machol eds., Amsterdam: North– Holland, 131–140.

Pollard GN and Pollard GH (2007). Optimal risk taking on first and second serves. In Proceedings of *Tennis Science & Technology 3*, S. Miller and J. Capel-Davies eds., London: International Tennis Federation, 273-280.

Pollard GN (2008). What is the best serving strategy? *J Med Sci Tennis* 13(2): 34-38.

Pollard GN, Pollard GH, Barnett T and Zeleznikow J (2009). Applying tennis match statistics to increase serving performance during a match in progress. Journal of Medicine and Science in Tennis 14(3), 16-19.

Tennis Data

Introduction

The men’s professional circuit has undergone several structural changes since the Open Era in 1968. The Association of Tennis Professionals or ATP was formed in September 1972. Since 1990, the association has organized the worldwide tennis tour for men and linked the title of the tour with the organization's name. In 1990 the organization was called the ATP Tour, which was renamed in 2001 as just ATP and the tour being called ATP Tour. In 2009 the name was changed again and is now known as the ATP World Tour. It is an evolution of the tour competitions previously known as Grand Prix tennis tournaments (1970 to 1989) and World Championship Tennis (1968 to 1990). The International Series (known before 2000 as the ATP World Series) and the International Series Gold were a series of professional tennis tournaments held internationally as part of the ATP from 1990 until 2008. The current structure being the ATP World Tour comprises a series of tournaments from ATP World Tour Finals, ATP World Tour Masters 1000, ATP World Tour 500 series, ATP World Tour 250 series and ATP Challenger Tour. Note that the ATP World Tour Masters 1000 had emerged from 1990; originally known as the ATP Championship Series Singles Week (1990-1993), Mercedes-Benz Super 9 (1993-1999), Tennis Masters Series (2000-2003), ATP Masters Series (2004-2008), and the present name of ATP World Tour Masters 1000 took effect in 2009. Note also that the ATP World Tour Finals was known as the ATP Tour World Championship (1990-1999) and the Tennis Masters Cup (2000-2008). The ITF Men's Circuit is a series of professional tennis tournaments held around the world that are organized by the International Tennis Federation. Originally, the ITF Men's Circuit consisted of Satellite tournaments, each of which took place over four weeks. However, in 1998, the ITF introduced Futures tournaments, allowing for greater flexibility in the organization of the tournaments for national associations, and participation in tournaments for players. Over time, the ratio of Futures tournaments to satellites increased until 2007, when satellites were eliminated entirely.

Similar tournament structural changes have occurred on the women’s professional circuit since the Open Era in 1968. The Women's Tennis Association (WTA), founded in 1973 by Billie Jean King, is the principal organizing body of women's professional tennis. Since 1983, the association has organized the worldwide tennis tour for women. Formed in 1970, the Virginia Slims Circuit eventually became the basis for the later named WTA Tour. The WTA Tier I, II, III and IV structure were a series of professional tennis tournaments held internationally as part of the WTA tour from 1988 until 2008. The current tournament structure of the WTA Tour was introduced in 2009 and comprises a series of tournaments from WTA Tour Championships, Premier, International and Challenger events. Note that the Premier tournaments consist of Premier Mandatory, Premier Five and Premier. The ITF Women's Circuit is a series of professional tennis tournaments run by the International Tennis Federation for female professional tennis players.

A vast amount of data is collected and stored in tennis either directly online or through various commercial software providers. This includes the typical set-by-set score line. For example Novak Djokovic defeated Andy Murray in the 2013 Australian Open final 6-7, 7-6, 6-3, 6-2. Match statistics may also be available after the completion of matches; particularly for grand slam events. However point-by-point data or match statistics (broken down by each set) are not as commonly available as the former. This article will outline for various data sources the coverage (by tournament type) and year commencing for each data type. To simplify the analysis the initialization of data for men's tennis will be from 1990; with the International Series tournament structure. Similarly for women's tennis the initialization of data will be from 1988; with the Tiered Series tournament structure. Summarizing data in this fashion could be useful for building decision support tools to enhance elite performance (Bedford et al, 2010). For example a coach may be interested in knowing a player's career average of points won on serve at a grand slam level as well as at an ATP World Tour Masters 1000 level. Whilst the focus is on the men's and women's professional singles circuit, similar methodology could be constructed for the men's and women's professional doubles circuits.

Method

Tournament classification

Table 1 provides the current tournament structure for the men's singles tour with the corresponding commencement year (with an initialization of 1990), number of tournaments played and winner's rating points in 2012. Note how the ATP World 500 series replaced the International Series Gold in 2009, and similar replacements occurred for the ATP World 250 series and ATP Challenger in 2009. Similarly, table 2 provides the current tournament structure for the women's singles tour with the corresponding commencement year (using an initialization of 1988), number of tournaments played and winner's rating points in 2012. Table 3 provides a comparison of tournament structures between the men's and women's professional tennis circuits. Note how the tournament structure from table 2 for the women's circuit is slightly modified to align with the men's professional circuit. This consists of splitting the ITF circuit between tournaments where the prize money for the tournament is 10K and >10K, combining Premier Mandatory and Premier 5 events, including the WTA Tournament of Champions in International events, and combining Challenger WTA 125s and ITF>10K events. The information from table 3 in tournament classification is used below in classifying data.

Category	**Year**	**Number of Tournaments (2012)**	**Winner's Rating Points (2012)**
Grand Slams	1990 onwards	4	2000
ATP World Tour Finals	1990 onwards	1	1100-1500
ATP World Tour Masters 1000	1990 onwards	9	1000
Olympic Games	1990 onwards	1 (every 4 years)	750

ATP World Tour 500 series International Series Gold	2009 onwards 1990-2008	11	500
ATP World Tour 250 series ATP International Series ATP World Series	2009 onwards 2000-2008 1990-1999	40	250
ATP Challenger Tour ATP Challenger Series	2009 onwards 1990-2008	148	80-125
Futures Satellites	1998 onwards 1990-2006	582	18-35

Table 1: Tournament structure for men's singles tour

Category	**Year**	**Number of Tournaments (2012)**	**Winner's Rating Points (2012)**
Grand Slams	1988 onwards	4	2000
WTA Tour Championships	1988 onwards	1	1050-1370
Premier Mandatory Tier I	2009 onwards 1988-2008	4	1000
Premier 5 Tier I	2009 onwards 1988-2008	6	900
Olympics	1988 onwards	1 (every 4 years)	685
Premier Tier II	2009 onwards 1988-2008	11	470
WTA Tournament of Champions	1998 onwards	1	366-435
International Tier III/Tier IV	2009 onwards 1988-2008	29	280
Challenger WTA 125s	2012 onwards	3	160
ITF Circuit	1988 onwards	487	12-150

Table 2: Tournament structure for women's singles tour

Category		**Number of Tournaments (2012)**		**Winner's Rating Points (2012)**	
Men	**Women**	**Men**	**Women**	**Men**	**Women**
Grand Slams	Grand Slams	4	4	2000	2000
ATP World Tour Finals	WTA Tour Championships	1	1	1100-1500	1050-1370
ATP World Tour Masters 1000	Premier Mandatory/Premier 5	9	10	1000	900-1000
Olympic Games	Olympics Games	1	1	750	685
ATP World Tour 500	Premier	11	11	500	470
ATP World Tour 250	International/WTA Tournament of Champions	40	30	250	280-435
ATP Challenger Tour	ITF>10K/Challenger WTA 125s	148	197	80-125	50-160

Futures	ITF 10K	582	293	18-35	12

Table 3: Comparison of tournament structures between the men's and women's tour

Data classification

Table 4 outlines data availability for the men's professional circuit. OnCourt[1] and Tennis Navigator[2] are commercially available software packages. The tennis ratings are given on a weekly basis for OnCourt (since 2003) and Tennis Navigator (since 2007). The Rating Tour refers to all the tournament types outlined in table 1. The Main Tour refers to tournaments in table 1 with the exclusion of the ATP Challenger Tour and Futures. OnCourt, Tennis Navigator and the ATP World Tour[3] provide set-by-set (incl. game score) score lines. However the ATP World Tour provides this information for all tournament types since 1990. OnCourt provides game-by-game score lines for the Main Tour from 2007 and provides point-by-point score lines for the Main Tour as well as the ATP Challenger Tour. Match statistics are provided for both OnCourt and Tennis Navigator. However Tennis Navigator provides match statistics broken down by each set for Grand Slam matches since 2003. The various Grand Slam sites also provide match statistics broken down by each set at the completion of matches. However the Grand Slam sites are more detailed than the standard match statistics given in Tennis Navigator (and OnCourt). For example Rally Stats are given for each set consisting of Approach Shots, Drop Shots, Ground Strokes, Lobs, Overhead Shots, Passing Shots and Volleys; and each shot categorized as a Forehand or Backhand Winner, a Forehand or Backhand Forced Error, or a Forehand or Backhand Unforced Error. Table 5 gives the Rally Statistics for the 1st Set of the 2013 Australian Open final between Novak Djokovic and Andy Murray. Table 6 is given similarly to table 4 for the women's professional circuit.

Type	**Type Breakdown**	**Source**	**Coverage**	**Year Commencing**
Ratings	Weekly	OnCourt	Rating Tour	2003
Ratings	Weekly	Tennis Navigator	Rating Tour	2007
Score line	Set-by-Set (incl. game score)	OnCourt	Main Tour ATP Challenger Tour Futures	1990 1998 2004
Score line	Set-by-Set (incl. game score)	Tennis Navigator	Main Tour ATP Challenger Tour	1980 2005
Score line	Set-by-Set (incl. game score)	ATP World Tour[3]	Main Tour ATP Challenger Tour Futures	1990 1990 1998
Score line	Game-by-Game	OnCourt	Main Tour	2007
Score line	Point-by-Point	OnCourt	Grand Slams Finals Masters 1000 Olympic Games	2009 2009 2009 2012

			500 series 250 series ATP Challenger Tour	2010 2010 2012
Match Statistics	Match	OnCourt	Grand Slams Finals Masters 1000 Olympic Games 500 series 250 series ATP Challenger Tour	2004 2006 2006 2012 2007 2007 2011
Match Statistics	Match	Tennis Navigator	Grand Slams* Finals Masters 1000 500 series 250 series	2003 2009 2009 2010 2010
Match Statistics	Set	Tennis Navigator	Grand Slams*	2003
Match Statistics	Set	Grand Slam sites	Grand Slams^	
Match Statistics	Set (incl. serve stats) (incl. return stats) (incl. rally stats) (incl. dir serve stats)	Grand Slam sites	Grand Slams*^	

Table 4: Data availability for men's professional circuit

* excludes qualifying matches

^ data available online for approximately one year

	Novak Djokovic						Andy Murray					
	Winners		Forced Errors		Unforced Errors		Winners		Forced Errors		Unforced Errors	
	FH	BH	FH	BH	FH	BH	FH	BH	FH	BH	FH	BH
Approach Shots	-	1	-	-	-	-	-	-	-	-	-	-
Drop Shots	-	1	-	-	-	-	-	-	-	-	-	1
Ground Strokes	1	2	2	3	11	10	4	-	1	2	3	7
Lobs	-	-	-	-	-	-	-	-	-	2	-	-
Overhead Shots	4	-	-	-	-	-	1	-	-	-	-	-
Passing Shots	1	1	-	-	-	-	-	-	2	1	-	-
Volleys	-	-	-	1	1	-	2	-	-	-	-	-

Table 5: Rally Statistics for the 1st Set of the 2013 Australian Open final between Novak Djokovic and Andy Murray

Type	**Type Breakdown**	**Source**	**Coverage**	**Year Commencing**
Ratings	Weekly	OnCourt	Rating Tour	2003
Ratings	Weekly	Tennis Navigator	Rating Tour	2004
Score line	Set-by-Set (incl. game score)	OnCourt	Main Tour Challenger (125s)/ITF>10K ITF 10K	1997 2002 2005
Score line	Set-by-Set (incl. game score)	Tennis Navigator	Main Tour	1995
Score line	Set-by-Set (incl. game score)	WTA[4]	Main Tour	1988
Score line	Game-by-Game	OnCourt	Main Tour	2007
Score line	Point-by-Point	OnCourt	Main Tour	2010
Match Statistics	Match	OnCourt	Grand Slams WTA Tour Championships Premier Mandatory/5 Olympics Premier International/ WTA Tournament of Champions	2004 2005 2006 2012 2006 2007
Match Statistics	Match	Tennis Navigator	Grand Slams*	2003
Match Statistics	Set	Tennis Navigator	Grand Slams*	2003
Match Statistics	Set	Grand Slam sites	Grand Slams^	
Match Statistics	Set (incl. serve stats) (incl. return stats) (incl. rally stats) (incl. dir serve stats)	Grand Slam sites	Grand Slams*^	

Table 6: Data availability for women's professional circuit

* excludes qualifying matches

^ data available online for approximately one year

Conclusions

Data availability for the women's and men's singles professional tennis circuits are given in concise tables by categorizing for the various data sources the coverage (tournament type) and year commencing. The results could be used in building decision support tools allowing for accessible data information for coaches in preparation for an upcoming match. Whilst the focus is on the men's and women's professional singles circuit, similar methodology could be constructed for the men's and women's professional doubles circuits.

References

Bedford A, Barnett T, Pollard GH and Pollard GN (2010). How the interpretation of match statistics affects player performance. Journal of Medicine and Science in Tennis 15(2), 23-27.

[1]OnCourt - www.oncourt.info
[2]Tennis Navigator - www.tennisnavigator.com
[3]ATP World Tour - www.atpworldtour.com/Scores/Archive-Event-Calendar.aspx
[4] WTA - www.wtatennis.com/tournament-archive

Resource Allocation

Introduction

In a sporting contest between two opposing sides, there are often strategic decisions to be made by the captain, coach or player, which could have an effect on the final outcome. Examples include when to take the batting power play in one-day cricket, when to pull the goalie in ice-hockey and how to optimize energy resources in a tennis match. A problem can be analyzed by assuming that one player/team makes all the decisions. More complex problems are where both players/teams are involved in the decision making process.

One area in which optimal allocation of energy is crucial is in contests with nested scoring systems (e.g. tennis). In these systems it is uncertain when the match will finish and players can win the match by winning well under half of the points played (Ferris 2003). Morris (1977), O'Donoghue (2001) and Pollard and Noble (2002) show that expending additional physical and mental effort on the important points in a game whilst relaxing on the unimportant points increases the chances of winning a game. In particular Morris (1977) states "*If he increased p from 0.60 to 0.61 on half his service points, and decreased from 0.60 to 0.59 on the unimportant half, he would increase his winning percentage by 0.0075 from 0.7357 to 0.7432*".

Other researchers have similarly addressed the question of optimal strategies to increase the probability of winning a tennis match. Barnett et al. (2004) concluded that an increase in effort is to be applied on every point of a game or every set of a match until either the game is finished or there are no increases remaining. Pollard and Pollard (2007) obtain optimal solutions for the maximum and minimum number of lifts available such that it is optimal not to lift and to lift respectively, for various point scores in a tiebreaker game and game scores in a tiebreaker set. Brimberg et al. (2004) model the situation where a player must allocate limited energy in a first-to-*n* match. They conclude that with only two possible energy choices for each game, it does not matter how energy is expended. With more than two possible energy choices, when the decision-maker falls behind in a match, s/he ought to switch to a more conservative approach by dividing her/his remaining energy evenly among all the possible remaining games.

The above references are modelled where only one player makes all the decisions in the match. This paper will account for the interactive nature of tennis and analyze a best-of-3 set match to include the situations where a) one player can apply an increase in effort on any set in the match, b) one player can vary effort about an overall mean, and where c) both players can apply an increase in effort on any set in the match. A conjecture is devised to obtain an optimal solution for a best-of-*N* set match, when both players can apply an increase in effort on any set in the match. The results from the analysis are used as a guide to strategy choices in sports where there are two opposing sides.

Analysis

Scenario a)

A best-of-3 set match is a contest where the first player to win 2 sets wins the match. Analyzing this system is non-trivial despite its relatively simple structure, because it is not certain that the third set will be played.

Consider the situation where a player has a constant probability p of winning a set. What is the probability of this player winning a best-of- 3 set match? The player can win the match by either winning in straight sets with probability p^2, losing the first set and winning the last two sets with probability $(1-p)p^2$ or winning the first and last set and losing the second set with probability $p(1-p)p$. Summing these, the probability of the player to win the match is given by $p^2(3-2p)$.

Now suppose a player increases his effort for one set at a match score in sets (e,f) (e= referred player's score, f=opponent's score), to change his probability of winning this set from p to $p+\varepsilon$, where $p+\varepsilon < 1$. This is equivalent to the opponent decreasing his effort at a match score (e,f) to change his probability of winning this set from $1-p$ to $1-p-\varepsilon$, since an increase of the probability of winning to one player is a decrease to the other player. On which set, should the player apply the increase to optimize their chances of winning the match? If the increase in effort is applied at (0,0), the probability for the player to win the match becomes $(p+\varepsilon)p(2-p) + (1-p-\varepsilon)p^2 = p^2(3-2p) + \varepsilon 2p(1-p)$. The same result is obtained if an increase in effort is applied at (1,1). Similarly the probabilities of a player to win the match when an increase in effort is applied at one of (1,0) or (0,1) is $p^2(3-2p) + \varepsilon p(1-p)$. Conditional on the match score reaching (1,0), the probability for a player to win the match when an increase in effort is applied at (1,0) or (1,1) is $p(2-p)+ \varepsilon(1-p)$; and conditional on the match score reaching (0,1), the chance for a player to win the match when an increase in effort is applied at (0,1) or (1,1) is $p^2+ \varepsilon p$. Table 1 gives the increase in probability when effort is applied throughout the match. The first set played begins with the match score at (0,0). The third set is played only if the match score reaches (1,1). The second set played occurs with the match score at either (1,0) or (0,1). The probability of a player winning the match when one increase in effort is applied on the first, second or third set played is equal to $p^2(3-2p) + \varepsilon 2p(1-p)$. Using this result and the results represented in Table 1, an increase in effort could be applied on any set played within the match, and the player has optimized their chances of winning.

Scenario b)

Now suppose a player adopts a strategy of increasing his effort on the first, second or third set played by ε, and decreases p on the first, second or third set played (but a different set played from that of the increase) by ε, where $0 < p+\varepsilon < 1$. Calculations show the chance of the player winning the match for this situation is equal to $p^2 (3-2p)+ \varepsilon^2(2p-1)$. When $p=½$, $2p-1 = 0$, and there is no change in the chances for either player to win the match. When $p> ½$, the chance for the player to win the match increases by $\varepsilon^2(2p\text{-}1)$ and therefore the opponent's chances to win the match decrease by $\varepsilon^2(2p\text{-}1)$. This implies that it is an advantage for the better player to vary his effort whilst maintaining his mean probability of winning a set. It follows by symmetry that the weaker player is disadvantaged by varying his effort.

Current match score	Match score at which an increase is applied	Increase in probability of winning match
(0,0)	(0,0)	$\varepsilon 2p(1\text{-}p)$
	(1,0)	$\varepsilon p(1\text{-}p)$
	(0,1)	$\varepsilon p(1\text{-}p)$
	(1,1)	$\varepsilon 2p(1\text{-}p)$
(1,0)	(1,0)	$\varepsilon(1\text{-}p)$
	(1,1)	$\varepsilon(1\text{-}p)$
(0,1)	(0,1)	εp
	(1,1)	εp
(1,1)	(1,1)	ε

Table 1. The increase in probability when effort is applied throughout the match.

Scenario c)

We now model the situation where both players can apply an increase in effort, which is represented by a two-person zero-sum game. For a best-of-3 set match, either player can apply an increase in effort at the first, second or third set played, resulting in a total of 9 possibilities. An increase in effort by ε at a set played from player A, results in increasing p to $p + \varepsilon$ ($p + \varepsilon < 1$), and an increase in effort by α at a set played from player B, results in decreasing p to $p - \alpha$ ($p - \alpha > 0$), where p represents the probability of player A winning a set. For the time being, it is assumed that both players must decide before the match has begun, on which set played that an increase is to be applied, and cannot change this choice throughout the match. Table 2 represents the probabilities of player A winning the match when an increase in effort is applied at the various sets played, where I_A and I_B represent an increase in effort at a set played by players A and B respectively. Notice that when both players apply an increase in effort on the same set played, the probability of player A winning the match is the same. Similarly, when both players apply an increase in effort on different sets played, the probability of player A winning the match is the same. When:

$$p^2(3 - 2p) + 2p(1 - p)(\varepsilon - \alpha) + \alpha\varepsilon(2p - 1) > p^2(3-2p) + 2p(1- p)(\varepsilon - \alpha)$$

$\rightarrow \alpha\varepsilon(2p\text{-}1) > 0$

$\rightarrow p > ½$

Similarly when:

$p^2(3 - 2p) + 2p(1 - p)(\varepsilon - \alpha) + \alpha\varepsilon(2p - 1) < p^2(3\text{-}2p) + 2p(1 - p)(\varepsilon - \alpha)$

$\rightarrow \alpha\varepsilon(2p\text{-}1) < 0$

$\rightarrow p < ½$

I_A	I_B	Probability of player A winning
0	0	$p^2(3\text{-}2p) + 2p(1- p)(\varepsilon - \alpha)$
1	0	$p^2(3 - 2p) + 2p(1 - p)(\varepsilon - \alpha) + \alpha\varepsilon(2p - 1)$
0	1	$p^2(3 - 2p) + 2p(1 - p)(\varepsilon - \alpha) + \alpha\varepsilon(2p - 1)$
1	1	$p^2(3\text{-}2p) + 2p(1- p)(\varepsilon - \alpha)$
2	0	$p^2(3 - 2p) + 2p(1 - p)(\varepsilon - \alpha) + \alpha\varepsilon(2p - 1)$
0	2	$p^2(3 - 2p) + 2p(1 - p)(\varepsilon - \alpha) + \alpha\varepsilon(2p - 1)$
1	2	$p^2(3 - 2p) + 2p(1 - p)(\varepsilon - \alpha) + \alpha\varepsilon(2p - 1)$
2	1	$p^2(3 - 2p) + 2p(1 - p)(\varepsilon - \alpha) + \alpha\varepsilon(2p - 1)$
2	2	$p^2(3\text{-}2p) + 2p(1- p)(\varepsilon - \alpha)$

Table 2. Probability of player A winning the match when an increase in effort is applied by both players at a set played in a match.

The increase in probability of winning for the better player when an increase in effort for both players is applied on different sets, is a result of the variability about the overall mean, as presented in the above section. Let $X = p^2(3\text{-}2p) + 2p(1- p)(\varepsilon - \alpha)$ and $Y = p^2(3 - 2p) + 2p(1 - p)(\varepsilon - \alpha) + \alpha\varepsilon(2p - 1)$. Let strategy *Ki* (*K:* {A,B}, *i :* {1, 2, 3}) refer to player *K* applying an increase in effort at *i* sets played. The game theory matrix is represented by:

	B1	B2	B3
A1	X	Y	Y
A2	Y	X	Y
A3	Y	Y	X

This matrix can easily be solved and the results indicate that players A and B should apply mixed strategies of A: (⅓, ⅓,⅓) and B: (⅓, ⅓,⅓). The value (v) of the game is then ⅓X + ⅔Y. For example if $p = 0.25$, $\varepsilon = \alpha = 0.1$, then $X = 0.15625$; $Y = 0.15125$ and $v = 0.1529$.

Suppose either player can now alter their strategies as the match is in progress. When should either player apply an increase in effort to optimize the usage of their available increase?

Consider the following analysis. Suppose at the start of the match, player A decides to apply an increase in effort at the first set played with a score line of (0,0), and player B decides to apply an increase in effort at the third set played with a score line (1,1). After the first set has been played, player B now has a decision to make on whether to stay with the initial strategy, by applying an increase in effort at the third set, or change strategies and apply an increase in effort at the second set played. As previously calculated in the above section, player B has the same probability of winning the match by applying an increase at the second or thirds sets played. Therefore player B could change their initial strategy by applying an increase in effort at the second set, and have optimized the usage of their available increase. Similarly, if player B decides at the start of the match to apply an increase in effort at the first set played, and player A decides to apply an increase in effort at the third set played, then player A could change their initial strategy by applying an increase in effort at the second set, and have optimized the usage of their available increase. This analysis is summarized as follows:

1. Both players are to apply an increase in effort at the first set played with probability of ⅓ .
2. If one player applies an increase in effort at the first set played, then the other player can decide to apply an increase in effort at either the second or third sets played. If neither player increased their effort at the first set played, then both players are to apply an increase in effort at the second set played with probability of ½
3. If the match reaches (1,1) and neither player has applied their increase in effort, then the increase in effort by both players must be applied at this state of the match.

Note that the mixed strategies of A: (⅓, ⅓,⅓) and B: (⅓, ⅓,⅓) still gives an optimal solution.

A best-of-3 set match is extended to a best-of-*N* set match in the case *N* odd by the following conjecture:

Suppose both players can apply one increase in effort in a best-of-N set match (N: odd integer). An increase in effort by ε at a set played from player A, results in increasing p to $p + \varepsilon$ ($p + \varepsilon < 1$), and an increase in effort by α at a set played from player B, results in decreasing p to $p - \alpha$ ($p - \alpha > 0$), where p represents the probability of player A winning a set. Then an optimal strategy for both players is to decide at the start of the match to apply the increase in effort with equal probability at N sets played, where the probability of applying the increase in effort at a set is given by $1/N$. The value of the game is given by $1/N\, X + (N\text{-}1)/N\, Y$, where X represents the probability of player A winning the match when an increase in effort by each player is applied at the same set played, and Y represents the probability of player A winning the match when an increase in effort by each player is applied at different sets played.

Applications

It was shown from Scenario b), where one player can vary effort around an overall mean, that it is advantageous for the better player to vary his effort whilst maintaining his mean probability of winning a set. It follows that the weaker player is disadvantaged by varying his effort. This finding is in line with those of Mulvey et. al. (2002), who modelled the distribution of scores in golf from a selection of players. Their results showed that the lowest mean scores do not necessarily imply a greater chance of winning, since the standard deviation of players' scores also contribute to the chances of winning. This finding also has other applications to sports, as variability can enhance performance. For example, tactical decisions need to be made during a middle or long distance running event on whether an athlete should break away from the leading pack early or to stay with the leading pack in an attempt for a sprint finish. An athlete may be better off to break away from the leading pack early as this could be their best chance of finishing in a reasonable position even though they possibly run the risk of finishing in the bottom positions by utilizing energy resources too early. This example attempts to show that an athlete may be better off by increasing their variability by taking levels of risk, even though their overall mean performance may be better by not taking such levels of risk.

It was shown from Scenario c), where both players can apply an increase in effort,- that when both players apply an increase in effort on a particular set in a best-of-N set match, mixed strategies for both players should be applied where the probability for each set is equally likely. It was shown from Scenario a), where one player can apply an increase in effort on any set in the match,-that applying this increase in effort on any set played within the match, optimizes the player's chances of winning. This also implies that a player could optimize their chances of winning by applying an increase in effort on each set played with a probability, where the probabilities of the increase for each set are equally likely. Since a solution to Scenario a) is equivalent to the solution for Scenario c), then a player can be no worse off by analyzing Scenario c) - where both players/teams are optimizing their energy resources.

Mixed strategy solutions in game theory, such as those applied here, assume that the probabilities are randomized and this could have applications in other sports. For example, the batting power play in one-day cricket, when to pull the goalie in ice-hockey and when to interchange players in football could be randomized. Tennis players are known to use an occasional change-up tactic on serve as a surprise factor. This could involve a serve-and-volley or a slower paced first serve with heavy topspin. Randomizing this decision process could improve performance, based on the expectations of the opponent.

Conclusions

It has been shown that when only one player can apply an increase in effort on any set in a best-of-3 set match, their chances of winning the match are the same irrespective of which set the increase is applied. However, variability about an overall mean for a best-of-3 set match gave an increased probability of winning the match for the better player. By analyzing a best-of-3 set match, where each player can apply one increase in effort, a mixed strategy solution was obtained, where an optimal strategy for each player was given by A: (⅓, ⅓, ⅓) and B: (⅓, ⅓, ⅓). A conjecture is devised for a two-person zero-sum to obtain an optimal solution for a best-of-*N* set match. Some applications are given to the theoretical results, which could be used by coaches and players to optimize performance by increasing variability and randomising decision events.

References

Barnett, T., Brown, A. and Clarke, S.R. (2004) Optimal Use of Tennis Resources, In Proceedings of the *7th Conference on Mathematics & Computers in Sport*, R.H. Morton and S. Ganesalingam eds., 57–65.

Brimberg, J., Hurley, W.J., and Lior, D.U. (2004) Allocating energy in a first-to-n match, *IMA Journal of Management Mathematics*, 15(1), 25–37.

Ferris, T.L.J. (2003) Emergence: An Illustration of the Concept for Education of Young Students, Proceedings of the *Thirteenth Annual International Symposium of the International Council on Systems Engineering*, Washington DC, 29 June-3 July, pp 945-956.

Morris, C. (1977) The Most Important Points in Tennis, *Optimal Strategies in Sports*, R.E. Machol (editor), 131-140.

Mulvey, J.M., Green, W. and Traub, C. (2002) Team Selection by Portfolio Optimization, *Science and Golf IV*, E. Thain (editor), Routledge, London, 305-316.

O'Donoghue, P. (2001) The Most Important Points in Grand Slam Singles Tennis, *Research Quarterly for Exercise and Sport No 2*, **72** (2001), 125-131.

Pollard, G. and Noble, K. (2002) The Effect of a Variation to the Assumption that the Probability of Winning a Point in Tennis is Constant, *6th Conference on Mathematics and Computers in Sport*, G.Cohen (editor), Bond University, Australia, 227-230.

Pollard, G.N. and Pollard, G.H. (2007) Importances 2: The most important points in a tiebreaker game, and the most important games in a set, In proceedings of *Tennis Science & Technology 3*, S. Miller and J. Capel-Davies (eds.), 293-300.