Sports Scheduling: Minimizing Travel for English Football Supporters

Graham Kendall and Stephan Westphal

Abstract The football authorities in England are responsible for generating the fixtures for the entire football season but the fixtures that are played over the Christmas period are given special consideration as they represent the minimum distances that are traveled by supporters when compared with fixtures played at other times of the year. The distances are minimized at this time of the year to save supporters having to travel long distances during the holiday period, which often coincides with periods of bad weather. In addition, the public transport system has limited services on some of the days in question. At this time of the year every team is required to play, which is not always the case for the rest of the season. When every team is required to play, we refer to this as a *complete fixture*. Additionally, each team has to to play a home game and an away game. Therefore, over the Christmas period we are required to produce two complete fixtures, where each team has to have a Home/Away pattern of HA or AH. In some seasons four complete fixtures are generated where each team is required to have a Home/Away pattern of HAHA (or AHAH). Whether two or four fixtures are generated there are various other constraints that have to be respected. For example, the same teams cannot play each other and we have to avoid (as far as possible) having some teams play at home on the same day. This chapter has three main elements. i) An analysis of seven seasons to classify them as two or four fixture seasons. ii) The presentation of a single mathematical model that is able to generate both two and four fixture schedules which adheres to all the required constraints. Additionally, the model is parameterized so that we can conduct a series of experiments. iii) Demonstrating that the model is able to produce solutions which are superior to the solutions that were used in practise (the *published fixtures*) and which are also superior to our previous work. The solutions we generate are

Graham Kendall
University of Nottingham, Nottingham, UK and University of Nottingham, Malaysia Campus, e-mail: graham.kendall@nottingham.edu.my or graham.kendall@nottingham.ac.uk

Stephan Westphal
Institute for Numerical and Applied Mathematics, Georg-August University, Germany e-mail: s.westphal@math.uni-goettingen.de

near optimal for the two fixture case. The four fixture case is more challenging and the solutions are about 16% of the lower bound. However, they are still a significant improvement on the fixtures that were actually used. We also show, through three experimental setups, that the problem owner might actually not want to accept the best solution with respect to the overall minimized distance but might want to take a slightly *worse* solution but which offers a guarantee as to the maximum distance that has to be traveled by the supporters within each division.

1 Introduction

In England, the football (soccer in the USA) league structure comprises four main divisions. These are generically called "FA Premiership" (20 teams), "FL Championship" (24 teams),"FL Championship" (24 teams) and "FL Championship 2" (24 teams). These names change with sponsorship arrangements for the given season. Within each division, a double round robin tournament is held, resulting in 2036 fixtures that have to be scheduled each season. Even though each division is an independent double round robin tournament, they cannot be scheduled in isolation from one another as there are a number constraints which operate across the divisions. For example, we should avoid, irrespective of which division they play in, certain teams playing at home on the same day (*pairing constraint*), only a certain number of FA Premiership teams based in London can play at home on the same day, only a total number of London based clubs (across all four divisions) can play at home on the same day and only a certain number of Manchester based clubs can play at home on the same day. These constraints are collective referred to as *geographical constraints*. All these constraints are captured in the model presented in Section 4.

When generating a schedule for the entire season, it is our belief that the football authorities initially schedule fixtures for the Christmas period. This means creating two or four sets of fixtures that will be used over two or four days. At this time of the year, every team is required to play (which is not always the case for other times in the season). We refer to such a schedule as a *complete fixture*. That is, a complete fixture ensures that all 92 teams play, representing 46 fixtures. Therefore, over the Christmas period we are required to generate either two or four complete fixtures. As well as respecting the pairing and geographical constraints, there are a number of further constraints that we have to respect over the Christmas period. For a two fixture schedule a team must play one game at home and one at an away venue (or away and then home); a so called home/away pattern of HA (resp. AH). For four complete fixtures the home away pattern must be HAHA (or AHAH). Furthermore, it is not permissible for teams to play each other twice over these two, or four, complete fixtures. For example, Chelsea cannot play Liverpool and later in the two or four sequence, Liverpool play Chelsea.

When generating these fixtures, the overall aim is to minimize the overall distance for all the clubs. Analyzing previous seasons (and personal correspondence with the football authorities) shows that this is indeed the primary objective of these fixtures. In this chapter, we are able to generate fixtures that are significant improvements over the published fixtures (i.e. those that were actually used) but we also present a number of experiments which indicates that the problem owner might prefer slightly worse solutions but which appear to be fairer to the clubs as it limits the maximum distance that a club would have to travel.

It is not clear why some seasons require two sets of fixtures to be generated, yet other seasons require four sets of fixtures to be generated. We thought that four fixtures were generated in order to complete the football season slightly earlier that usual to enable the national side more time to prepare for a Summer tournament (the FIFA World Cup or the UEFA European Championship). However, the data does not support this view (see Section 3). However, due to the methodology proposed in this chapter, the football authorities could easily generate both two or four completes fixtures and decide which one they prefer.

To assist other researchers we note that all the published fixtures were obtained from the Rothmans/Sky Sports Yearbooks Rollin and Rollin (2002, 2003, 2004, 2005, 2006, 2007, 2008). The distance information was collated by ourselves using one of the UK motoring web sites where we entered the to/from postcodes of the football clubs to get the driving distance between the clubs. This, we believe is preferable to using the straight line distance. As the driving distances will change over time, we have made these distances available at (*for reviewers: we will make the data available on or our own web site*).

This chapter is organised as follows. In the next section we provide some background to sports scheduling. In Section 3 we analyze the previous season's fixtures to try and ascertain when it is required to generate two or four complete fixtures. The analysis is inconclusive but we believe that it is interesting to present this data for future researchers. In Section 4 we present our mathematical model, which is capable of generating two or four complete fixtures. In Section 5 we describe the various experiments that we conduct, followed by the results for each experiment. We discuss the results in Section 6 and conclude the chapter in Section 7.

2 Background

Various algorithms exist which produce double round robin tournaments, with the most well known probably being the polygon construction method (Dinitz et al (2006)). We are unable to use this method, in its raw form, as the generated fixtures would not be acceptable to all interested parties. That is, it would generate a valid double round robin tournament but the schedule would not adhere to other constraints imposed by the football clubs, football authorities, the supporters, the police etc. Nor would it minimize the distances, which is the prime objective.

Previous work has considered the minimization of travel distances for sports schedules. Costa Costa (1995), for example, investigated the scheduling requirements of the National Hockey League, where one of the factors was to minimize the distances traveled. Recent work Westphal (2011) has investigated reducing the distances that have to be driven on 2nd January 2012 for the German Basketball League. The fixtures were such that they form a minimum weight perfect matching (with respect to distances). This provides evidence that this area of sports scheduling is important even for relatively small leagues, and even when only one day is involved. The introduction of the Traveling Tournament Problem Easton et al (2001), using distances based on road trips that have to be undertaken by Major League Baseball teams in the United States, has helped promote research interest in this area. See, for example Crauwels and van Oudheusden (2002); Ribeiro and Urrutia (2004); Easton et al (2003); Westphal and Noparlik (2010), with the best results being reported in Anagnostopoulos et al (2006). An up to date list of the best known solutions, as well as details of all the instances, can be found at Trick (2009).

Urrutia and Ribeiro Urrutia and Ribeiro (2004) have shown that minimizing distance and maximizing breaks (two consecutive home or away games) is equivalent. This followed previous work de Werra (1981, 1988); Elf et al (2003) showing how to construct schedules with the minimum number of breaks.

Overviews and surveys of sports scheduling can be found in Easton et al (2004); Knust (2009); Rasmussen and Trick (2008); Kendall et al (2010a).

The problem that we consider in this chapter is the minimization of the distance traveled for two (or four) complete fixtures. These two (or four) complete fixtures can be used over the Christmas period when, for a variety of reasons, teams wish to limit the amount of traveling. Note, that this is a different problem to the Traveling Tournament Problem (Easton et al (2003)), which assumes that teams go on road trips, and so the total distance traveled over a season can be minimized. In English football, there is no concept of road trips, so the overall distance cannot be minimized. However, we are able to minimize the distance on certain days. Kendall Kendall (2008) adopted a two-phase approach to produce two complete fixtures with minimal distances. A depth first search was used to produce a complete fixture for one day, for each division. A further depth first search created another set of fixtures for another day. This process produced eight separate fixtures which adhered to some of the constraints (e.g. a team plays at home on one day and away on the other) but had not yet addressed the constraints with respect to pair clashes (where certain teams cannot play at home on the same day, see Appendix C and Table A2 in Kendall (2008)), the number of teams playing in London etc. (see Appendix D in Kendall (2008)). The fixture lists from the depth first searches were input to a local search procedure which aimed to satisfy the remaining constraints, whilst minimizing the overall distance traveled. The output of the local search, and a post-process operation to ensure feasibility, produced the results in Table 32.

3 Fixture Analysis

In Kendall (2008) an analysis was given of the four seasons considered in that paper. In this section we provide a more comprehensive analysis as we are now considering three additional seasons and we also extend the analysis to include four fixtures. For each season we consider the fixtures that were played around the Christmas period, seeking to find home and away patterns that we can use to classify it as a two or a four fixture. We also look at the distances and state whether the distances traveled for these fixtures are the minimum when compared to other complete fixtures in the season. We end up with a classification for each season.

3.1 Season 2002-2003

This season has four sets of complete fixtures (see Table 1) around the Christmas/New Year period. The fixtures played on 26th December and 1st January represent the lowest distances of any complete fixtures throughout the season. They also exhibit the property that if a team plays home on one day, they play away on the other (and vice versa) (i.e. HA or AH). The other complete fixtures (20/21/22/23 Dec and 28/29 Dec) are significantly higher with respect to distances, and there are no other complete fixtures in the season that have lower distances. In addition, the four complete fixtures do NOT have a HAHA (resp. AHAH) sequence for home and away patterns for each team. Therefore, this season is classified as a two fixture season, with a total of (3820+3964)=7784.

Table 1 Candidate complete fixtures for the 2002-2003 season. The selected fixtures are in bold and this season is classified as a *two* fixture season (see text for details)

Dates	# of fixtures	Distance
20th Dec 2002	4	484
21st Dec 2002	40	6016
22nd Dec 2002	1	1
23rd Dec 2002	1	199
	Total	6700
26th Dec 2002	46	3820
	Total	3820
28th Dec 2002	43	6871
29th Dec 2002	3	712
	Total	7583
1st Jan 2003	46	3964
	Total	3964

3.2 Season 2003-2004

This season has three sets of complete fixtures (see Table 2) around the Christmas/New Year period. The fixtures played on 26th and 28th December represent the lowest distances of any complete fixtures throughout the season. They also exhibit the property that if a team plays at home on one day, they play away on the other (and vice versa) (i.e. HA or AH). The other complete fixture (20th December) is higher with respect to distances, and there are other complete fixtures in the season that have lower distances. Therefore, this season is classified as a two fixture season, with a total of (3837+4342)=8179.

Table 2 Candidate complete fixtures for the 2003-2004 season. The selected fixtures are in bold and this season is classified as a *two* fixture season (see text for details)

Dates	# of fixtures	Distance
20th Dec 2003	46	6295
	Total	6295
26th Dec 2003	46	3837
	Total	3837
28th Dec 2003	46	4342
	Total	4342

3.3 Season 2004-2005

This season has five sets of complete fixtures (see Table 3) around the Christmas/New Year period. The fixtures played on 26th December are the lowest distances of any complete fixtures throughout the entire season. The fixtures on the 28th/29th are also amongst the minimal distances. There are some lower distances (e.g. 11th-13th September, 4985; 5th March, 5852; 23rd April, 5813) but we have to bear in mind that the fixtures on 26th December and 28th/29th December adhere to the HA (resp. AH) constraint. The fixtures on the 1st and 3rd Jan, although not being the lowest distances in the season for complete fixtures, do adhere to the HA (resp. AH) constraint. The fixtures on the 18th-20th Dec can be ignored as they do not have a HA (resp. AH) relationship with any of the other fixtures. Therefore, this season is classified as a four fixture season, with four sets of complete fixtures (4563+6449=11,012 and 5122+7139=12,261), giving a total of (11,012+12,261)=23,273). When we later analyze this season (see Section 6.3.2) as a two fixture schedule we use 26th December (4563) and 28th/29th December (6449) (total of 11012) as the comparator as these follow a HA (resp. AH) pattern and these are the lowest distances from the two sets of complete fixtures.

Table 3 Candidate complete fixtures for the 2004-2005 season. The selected fixtures are in bold and this season is classified as a *four* fixture season (see text for details)

Dates	# of fixtures	Distance
18th Dec 2004	44	5758
19th Dec 2004	1	79
20th Dec 2004	1	15
	Total	5852
26th Dec 2004	46	4563
	Total	4563
28th Dec 2004	45	6164
29th Dec 2004	1	285
	Total	6449
1st Jan 2005	46	5122
	Total	5122
3rd Jan 2005	46	7139
	Total	7139

3.4 Season 2005-2006

This season has four sets of complete fixtures (see Table 4) around the Christmas/New Year period. The fixtures are amongst the lowest across the entire season. There are some equally low distances, however, those on the 17th April and 1st April are a reverse of those on 26th December and 31st December resp., and so could not be used over Christmas as it would violate the no reverse constraint. The four sets of fixture adhere to the HAHA (resp. AHAH) constraint. Therefore, this season is classified as a four fixture season, with two sets of complete fixtures (4295+6331=10,626 and 4488+6645=11,333), giving a total of (10,626+11,333)=21,959 for the four complete fixtures. When we analyze the two fixture case (see Section 6.2), we use the 26th/28th December as these are the minimum of the two sets of complete fixtures.

3.5 Season 2006-2007

This season has four sets of complete fixtures (see Table 5) around the Christmas/New Year period. Although each team plays four complete fixtures, the home/away patterns are HAAH (resp. AHHA), rather than the more usual HAHA (resp. AHAH). However, we have still classified this season as a four fixture season, with two sets of complete fixtures (7904+3857=11,761 and 7324+4582=11,906), giving a total of (11,761+11,906)=23,667 for the four complete fixtures. When we later analyze this season as a two fixture schedule (see Section 6.3.3) we will 26th/27th December 2006 (3857) and 1st January 2007 (4582) (total of 8439) as the comparator as these

Table 4 Candidate complete fixtures for the 2005-2006 season. The selected fixtures are in bold and this season is classified as a *four* fixture season (see text for details)

Dates	# of fixtures	Distance
26th Dec 2005	46	4295
	Total	4295
28th Dec 2005	46	6331
	Total	6331
31st Dec 2006	46	4488
	Total	4488
2nd Jan 2006	45	6648
3rd Jan 2006	1	197
	Total	6845

follow a HA (resp. AH) pattern and these are the two lowest distances, so it is a fairer comparison.

Table 5 Candidate complete fixtures for the 2006-2007 season. The selected fixtures are in bold and this season is classified as a *four* fixture season (see text for details)

Dates	# of fixtures	Distance
23rd Dec 2006	46	7904
	Total	7904
26th Dec 2006	45	3843
27th Dec 2006	1	14
	Total	3857
30th Dec 2006	46	7324
	Total	7324
1st Jan 2007	46	4582
	Total	4582

3.6 Season 2007-2008

This season has four sets of complete fixtures (see Table 6) around the Christmas/New Year period. Like 2006/2007 the home away patterns follow HAAH (resp. AHHA), rather than HAHA (resp. AHAH). However, we still classify this season as a four fixture season, with two sets of complete fixtures (6943+4459=11,402 and 7226+4085=11,311), giving a total of (11,402+11,311)=22,713 for the four complete fixtures. When we later analyze this season as a two fixture schedule (see Section 6.3.4) we will 26th December 2007 (4459) and 1st/2nd January 2008 (4085) (total of 8544) as the comparator as these follow a HA (resp. AH) pattern and these are the two lowest distances, so it is a fairer comparison.

Table 6 Candidate complete fixtures for the 2007-2008 season. The selected fixtures are in bold and this season is classified as a *four* fixture season (see text for details)

Dates	# of fixtures	Distance
21st Dec 2007	4	276
22nd Dec 2007	42	6667
	Total	6943
26th Dec 2007	46	4459
	Total	4459
29th Dec 2007	46	7226
	Total	7226
1st Jan 2008	45	3991
2nd Jan 2008	1	94
	Total	4085

3.7 Season 2008-2009

This season has three sets of complete fixtures (see Table 7) around the Christmas/New Year period. The three sets of fixture follow a HAH (resp. AHA) pattern. However, the fixtures on the 26th and 28th December are the lowest distances and we use those fixtures and classify the season as a two fixture season, with two complete fixtures (4548+4764=9,312).

Table 7 Candidate complete fixtures for the 2008-2009 season. The selected fixtures are in bold and this season is classified as a *two* fixture season (see text for details)

Dates	# of fixtures	Distance
20th Dec 2008	46	7709
	Total	7709
26th Dec 2008	46	4548
	Total	4548
28th Dec 2008	46	4764
	Total	4764

3.8 Discussion

Of the seven seasons that we study in this chapter, three of them are classified as *two* fixture seasons, with the other four being classified as *four* fixture seasons (see Table 8). We initially believed that the reason a season was classified as a four fixture season was because the football authorities wanted the season to end slightly early to enable the national team to train together in preparation for the tournament. However, this appears not to be the case as we would have expected seasons 2003-2004, 2005-2006 and 2007-2008 to be classified as four fixture seasons and to finish earlier than the other seasons (at least with respect to the Premier division). The data does not support this assumption and we are unsure why some seasons have four complete fixtures at Christmas, and others have two.

When we carry out our experiments, we treat each season as both two and four fixture season so that other researchers have the data for comparative purposes and also to demonstrate that we are able to generate both type of fixtures for the seven seasons that we study.

Table 8 This table shows whether the football authorities generated a two or four fixture schedule over the holiday period. In all cases, these fixtures represent the minimum distances between clubs when compared against fixtures that are used at other times in the season. We also show whether the season was a World Cup or European Championship year.

Season	Two or Four	End Date (Prem)	End Date (Others)	World or Euro?
2002-2003	Two	11th May 2003	4th May 2003	
2003-2004	Two	15th May 2004	9th May 2004	Euro
2004-2005	Four	14th May 2005	8th May 2005	
2005-2006	Four	7th May 2006	6th May 2006	World
2006-2007	Four	13th May 2007	6th May 2007	
2007-2008	Four	11th May 2008	4th May 2008	Euro
2008-2009	Two	28th May 2009	3rd May 2009	

4 Mathematical Model

In earlier work we presented a naive approach Kendall (2008), and a slightly more sophisticated approach Kendall et al (2010b), in order to tackle the problem addressed in this chapter. These previous works had shortcomings, which are addressed here. Firstly, we only generated two complete fixtures, with the generation of four fixtures being left as future work. Secondly, both previous approaches used a two phase methodology. In the first phase fixtures were generated for individual divisions, without taking into account any constraints that operated across division boundaries. In the second phase, a local search was utilized that removed any hard constraints that were present and also minimized the soft constraint violations. The previous approaches could be time consuming. In particular, Kendall (2008), took upwards of 20 hours for the depth first search phase. Finally, the previous approaches utilized meta-heuristics and so the solutions were not provably optimal. Indeed, the results presented in this chapter are superior to our previous work which had already improved on the published fixtures. For reference, our previous results are summarized in Appendix A.

In this chapter we address these issues by presenting a mathematical formulation that attempts to solve the model in a single phase. That is, we consider all four divisions, eliminating the need for a local search phase to resolve hard constraint violations as the minimization of soft constraint violations.

The model is as follows, with explanations after:

Indices

L	the set of leagues
T	the set of teams
T_l	the set of teams belonging to league l
H	the set of days $\{1,2,\ldots,k\}$
P	the set of paired teams
R	the set of divisions

Decision Variables

$x_{i,j,d}$	1 if team i is playing team j on day d at i's site
$h_{i,d}$	1 if i is playing at home on day d
$y_{i,j,d}$	1 if the paired teams i and j play both at home on day d

Parameters

$D_{i,j}$	the distance (in miles) between team i and team j
L_i	1 if team i is a London-based club
M_i	1 if team i is a Greater Manchester-based club
Q_i	1 if team i is a Premier club
β_l	The maximum number of clubs based in London which can play at home on the same day. $\beta_l = 6$.
β_m	The maximum number of clubs based in Greater Manchester which can play at home on the same day. $\beta_m = 4$.
β_q	The maximum number of Premier Division clubs based in London which can play at home on the same day. $\beta_q = 3$.
δ_r	The maximum allowed travel distance for teams in division r.
γ	The maximum number of allowed pair clashes.

Objective Function

The objective function minimizes the total distance by all the teams and furthermore helps to adjust the *y*-variables correctly.

$$\min \sum_{h\in H, l\in L, i,j\in T_l} D_{i,j}\cdot x_{i,j,h} + \sum_{h\in H, \{i,j\}\in P} 0.01\cdot y_{i,j,h} \tag{1}$$

Subject to

Every team plays exactly one match per day.

$$\sum_{j\in T_l\setminus\{i\}} (x_{i,j,d}+x_{j,i,d}) = 1 \qquad \forall\, l\in L,\ i\in T_l,\ d\in H \tag{2}$$

Every pair of teams meets each other at most once.

$$\sum_{d\in H} \left(x_{i,j,d}+x_{j,i,d}\right) \le 1 \qquad \forall\, l\in L,\ i,j\in T_l \tag{3}$$

Paired teams are not allowed to play against each other.

$$x_{i,j,d} = 0 \qquad \forall\, d\in H, l\in L, \{i,j\}\in (T_l\times T_l)\cap P \tag{4}$$

Teams in division r are not allowed to travel a distance greater than δ_r miles.

$$x_{i,j,d} = 0 \qquad \forall\, d\in H, l\in L, i,j\in T_l : D_{i,j} > \delta_r \tag{5}$$

The following constraints couple fixture variables x to home-variables h.

$$\sum_{j\in T_l\setminus\{i\}} x_{i,j,d} = h_{i,d} \qquad \forall\, l\in L,\ i\in T_l,\ d\in H \tag{6}$$

Every team plays exactly one home game in two successive days (which implies exactly one away every in those two days)

$$h_{i,d}+h_{i,d+1} = 1 \qquad \forall\, l\in L,\ i\in T_l,\ d\in H\setminus\{k\} \tag{7}$$

Together with the objective function the following inequality ensures that $y_{i,j}=1$ if and only if the paired teams i and j play both at home on day d.

$$h_{i,d}+h_{j,d} \le 1+y_{i,j,d} \qquad \forall\, d\in H, \{i,j\}\in P \tag{8}$$

There are not more than γ pair clashes.

$$\sum_{d\in H, \{i,j\}\in P} y_{i,j,d} \le \gamma \tag{9}$$

The maximum number of Greater Manchester-based clubs playing at home on any of the holidays must not exceed a certain threshold.

$$\sum_{i \in T} L_i \cdot h_{i,d} \leq \beta_l \qquad \forall\, d \in H \qquad (10)$$

The maximum number of London-based clubs playing at home on any of the holidays must not exceed a certain threshold.

$$\sum_{i \in T} Q_i \cdot h_{i,d} \leq \beta_q \qquad \forall\, d \in H \qquad (11)$$

The maximum number of London-based Premier Division clubs playing at home on any of the holidays must not exceed a certain threshold.

$$\sum_{i \in T} M_i \cdot h_{i,d} \leq \beta_m \qquad \forall\, d \in H \qquad (12)$$

Notes

1. $L = \{$*FA Premeirship*, *FL Championship*, *FL Championship 1*, *FL Championship 2*$\}$. These are the four main divisions in the English league. The names of the divisions change in line with sponsorship agreements.
2. $|T|$ = 92, these being made up from 20 teams in the Premier division and 24 teams in each of the other three divisions.
3. H is the set of days, which will either be $\{1,2\}$ when generating a two fixture schedule or $\{1,2,3,4\}$ when generating a four fixture schedule.
4. P = a set of teams that are paired. If two teams are *paired* they, ideally, should not play at home on the same day. However, it is impossible to have zero pairing violations so we allow the same number that were present in the published fixtures (see Table 33 for the number of pairing violations that we allow). Details of the actual paired teams are given in Kendall (2008). They are not reproduced here for reasons of space.
5. γ defines the number of pair clashes that we allow. In the model, γ takes different values for each season (see Table 33). In previous work we defined separate values for Boxing Day and News Years Day (see Kendall (2008)) but this is no longer valid as we are producing schedules for both these days and also for four days. However the values used in Kendall (2008) are still used but are added together.
6. δ_r defines the maximum distance that can be traveled by any single team in division r. In our experiments, we try different values for δ_r, to test its effect.
7. Equation 1 minimizes the overall distance. The second term ensures that $y_{i,j,h} = 0$ if the paired teams i and j do not both play at home on day d.

5 Experimental Setup and Results

Using the model from Section 4 we used CPLEX 12.2 in order to solve various instances of the model so that we could explore several scenarios. All experiments were run on an Acer Ferrari 1100 laptop, with a 2.29 GHz processor (AMD Turion 64X2 Mobile, Technology TL-66), with 2.29GB of RAM and running Windows XP Professional (Version 2002, SP3)). We allowed CPLEX to run for 300 seconds (five minutes) and 7200 seconds (2 hours) for each experiment. Each scenario is presented below, along with the results. For each scenario we run two experiments $H = \{1,2\}$ (to capture the two day case) and $H = \{1,2,3,4\}$ (to capture the four day case).

We note that for all the solutions we present, they are an improvement on the published fixtures, as well as being an improvement on our previous work (see Table 32 in Appendix A).

In presenting the results, if the *Gap* is less than 0.01% (the default termination criteria for CPLEX), this indicates that a near optimal solution was found before the time expired. If the value in the *Seconds* column is less than 7200 (and the gap is greater than 0.01%), it indicates that CPLEX ran out of memory at the time shown and the result reported is the incumbent solution at that time. We do not report the time for the 300 second experiments as CPLEX never reported an out of memory condition and, unless a value of 0.01% is reported, it ran for the full 300 seconds.

5.1 Experiment 1: $\delta_r = \infty$

This experiment sets no limit on the distance that teams are allowed to travel. This, provides the most flexibility, as any team can play against any other team. The potential drawback is that some teams may travel far greater distances than others, although the total overall distance might be suitably minimized. In these experiments $\delta_0=\delta_1=\delta_2=\delta_3 = \infty$.

5.2 Results for Experiment 1: $\delta_r = \infty$

Allowing the algorithm to run for 300 seconds for the two day case (Table 9) we are able to find near optimal solutions. Allowing the algorithm to run for 7200 seconds (2 hours) we are able to further improve on the solutions generated. Season 2002-2003 was the only one where we could not get within 1% of the optimal solution.

For the four day case (Table 10) we are able to find solutions which are typically around 15% of the lower bound, although it seems appropriate to allow the algorithm to run for 7200 seconds, as this does provide better solutions that just allowing 300 seconds. There is further work to be done to decrease the optimality gap even further. However we note that for the four seasons that are classified as four fixture seasons (see Table 8) that we are able to produce much better solutions than the published fixtures (compare with see Table 32).

Table 9 Results: Experiment 1a: $\delta_r = \infty$ (two day case: $H = \{1,2\}$)

	300 Seconds			7200 Seconds			
Season	**LB**	**Found**	**% Gap**	**LB**	**Found**	**% Gap**	**Seconds**
2002-2003	4556.49	4905.09	7.11	4692.07	4801.09	2.27	5640
2003-2004	5172.79	5225.11	1.00	5185.91	5209.11	0.45	
2004-2005	5107.88	5182.10	1.43	5134.64	5161.10	0.51	4065
2005-2006	5037.63	5038.13	0.01	5037.63	5038.13	0.01	
2006-2007	5271.32	5373.11	1.89	5294.53	5308.11	0.26	
2007-2008	5002.78	5043.12	0.80	5019.79	5034.12	0.28	
2008-2009	5212.14	5245.10	0.63	5243.42	5244.10	0.01	3624

Table 10 Results: Experiment 1b: $\delta_r = \infty$ (four day case: $H = \{1,2,3,4\}$)

	300 Seconds			7200 Seconds			
Season	**LB**	**Found**	**% Gap**	**LB**	**Found**	**% Gap**	**Seconds**
2002-2003	11368.16	13995.18	18.77	11376.85	13813.18	17.64	
2003-2004	11890.66	14288.22	16.78	11896.10	13966.22	14.82	
2004-2005	12036.40	14255.20	15.56	12039.56	13605.20	11.51	
2005-2006	12221.16	14075.26	13.17	12221.19	13785.26	11.35	
2006-2007	12388.06	14706.22	15.76	12408.73	14262.22	13.00	
2007-2008	11982.16	14971.22	19.97	11984.65	14089.24	14.94	
2008-2009	12264.46	19015.14	35.50	12282.52	14671.20	16.28	

5.3 *Experiment 2: δ_r = maximum*

A potential problem with experiment 1 is that some teams may have to travel large distances so that others can travel shorter instances. In this experiment we set a *global* maximum distance such that no team can exceed that distance. If this value is too restrictive there will not be any feasible solutions. To give an example. In the 2003-2004 season, Plymouth's distances from the other teams in its division are (in ascending order) {119, 134, 162, 210, 214, 216, 219, 231, 246, 248, 254, 276, 277, 282, 284, 290, 292, 296, 300, 303, 321, 347, 389}. By simple inspection we can see that if set the maximum travel distance too low (in this case below 210) then it is impossible to generate a four fixture schedule as Plymouth will not be able to play four fixtures. Plymouth is often the team that will define the maximum travel distance, but it may not always be the case. By inspecting each season we can set the maximum distance, both for two season fixtures and for four season fixtures.

These are presented in Table 11. It should be noted that there is just one value for each season. That is, δ_0=δ_1=δ_2=δ_3. To continue the example from above, Plymouth must travel 210 miles, in the 2003-2004 season, and this is the value that is applied to every team, *in every division*. This means that teams in the other leagues can also travel up to this maximum distance.

Table 11 Maximum distances for each season that will enable a feasible schedule to be generated. These values define the maximum distance that at least one team has to travel and we set this as a maximum distance that all teams are able to travel. With reference to the model we set δ_r to the value in each cell depending on the season and whether we are generating a two or a four fixture schedule. For each experiment δ_0=δ_1=δ_2=δ_3.

Season	Two	Four
2002-2003	153	165
2003-2004	134	210
2004-2005	199	214
2005-2006	160	209
2006-2007	154	212
2007-2008	153	202
2008-2009	153	190

One potential drawback with this approach is that we are effectively dictating the fixtures for certain teams. Plymouth (in 2003-2004) will be forced to play Bournemouth (134 miles), Brentford (210 miles), Bristol City (119 miles), Swindon Town (162 miles). For the two fixture season (with a maximum distance of 134), Plymouth will be forced to play Bournemouth (134 miles), Bristol City (119 miles).

Another potential drawback is that we are giving too much scope to other divisions, as they are allowed to use the same maximum traveling distance as the division which has imposed the upper limit. We consider an extension of the model in Section 5.5 by setting δ_r for each division.

5.4 Results for Experiment 2: δ_r = maximum

Table 12 presents the results for the two day case. Like experiment 1 we are able to produce results to within 1% of optimality, with the exception of season 2002-2003.

For the four day case (Table 13), the 2002-2003 season proved to be intractable in that we never found an incumbent solution even after 7200 seconds. The other seasons produced similar results to the first experiment (i.e. around 15% of the lower bound).

Table 12 Results: Experiment 2a: δ_r **=** ***maximum*** **(two day case:** $H = \{1,2\}$**)**

	300 Seconds			7200 Seconds			
Season	**LB**	**Found**	**% Gap**	**LB**	**Found**	**% Gap**	**Seconds**
2002-2003	4742.30	4862.09	2.46	4767.61	4862.09	2.54	1800
2003-2004	5284.99	5311.11	0.49	5295.95	5311.11	0.29	
2004-2005	5176.87	5212.10	0.68	5199.59	5212.10	0.24	
2005-2006	5039.63	5040.13	0.01	5039.63	5040.13	0.01	
2006-2007	5305.31	5358.11	0.99	5323.76	5358.11	0.64	
2007-2008	5084.33	5096.12	0.23	5095.62	5096.12	0.01	
2008-2009	5337.95	5365.10	0.51	5364.56	5365.10	0.01	

Table 13 Results: Experiment 2b: δ_r **=** ***maximum*** **(four day case:** $H = \{1,2,3,4\}$**)**

	300 Seconds			7200 Seconds			
Season	**LB**	**Found**	**% Gap**	**LB**	**Found**	**% Gap**	**Seconds**
2002-2003	11394.16	-	-	11404.53	-	-	
2003-2004	11984.64	14825.22	19.16	11990.19	13805.22	13.15	
2004-2005	12088.16	14614.20	17.28	12089.37	13782.20	12.28	
2005-2006	12237.16	18128.22	32.50	12237.80	13884.26	11.86	
2006-2007	12439.16	15004.22	17.10	12445.41	14619.22	14.87	
2007-2008	12007.80	14938.24	19.62	12010.51	14011.24	14.28	
2008-2009	12427.14	-	-	12469.91	14962.20	16.66	

5.5 *Experiment 3:* δ_r *= maximum for each division*

In experiment 2, a global δ_r (i.e. δ_0=δ_1=δ_2=δ_3) was used across all divisions. In this experiment we explore if having a δ_r for each division is beneficial. We did plan to derive δ_r in the same way as experiment 2. However, it cannot easily be done by inspection, An example will explain why. Consider the 2005-2006 season, Premier Division. If we look for the team that has to travel the furthest, we find that Portsmouth has the following travel distances (in ascending order); $\{70, 71, 79, 79, 82, 82, 153, 161, 165, 243, 245, 247, 255, 256, 257, 268, 311, 336, 341\}$. Portsmouth has to travel at least 71 miles to complete a two fixture schedule (and 79 miles for a four fixture schedule). Therefore we can set δ_0=71. However, if we analyze this, we can see that this will mean that Portsmouth will play Fulham (70 miles away) and Chelsea (71 miles away). Looking at other teams, we note that Newcastle also only has two potential fixtures (Sunderland 15 miles away and Middlesbrough 45 miles away). The other fixtures for Newcastle are all greater than 71 miles. The problem is, Newcastle cannot play Sunderland, as they are paired. Therefore, if we set δ_0=71, there is no feasible solution. However, this is a simple case and it is not always obvious what values should be used, especially when we look at the four fixture case. We could use something such as constraint programming to determine suitable values but as we already have a model we decided to use that. Therefore, in order to derive the value for each division we proceeded as follows:

1. For each team in the Premiership, obtain their distance vectors (as we did for Portsmouth above) and sort each one in ascending order.
2. Find the team for the Premiership that has to travel the largest distance in the second position of the sorted distance vectors. In the example above, this will be Portsmouth whose distance vector is {70, 71, 79, 79, 82, 82, 153, 161, 165, 243, 245, 247, 255, 256, 257, 268, 311, 336, 341}. Element two is 71, which is the largest value for all teams.
3. Find the team that has the largest distance in the third position of the sorted distance vectors. In our 2005-2006 example, this is Newcastle {15, 45, 128, 144, 152, 153, 155, 170, 171, 200, 203, 212, 279, 280, 280, 285, 286, 287, 341}, so we take the value of 128.
4. We continue this process, taking the maximum values from the fourth, fifth, sixth values etc. from the sorted distance vectors. We do not need to carry out a complete analysis (although it is not time consuming, we simply used the SMALL function in Excel) as we do not need all the values.
5. This leads to a vector of {71, 128, 144, 152, ...}.
6. We now solve the model using $\delta_0 = 71$ and δ_r=∞ for all the other divisions (i.e. $\delta_1=\delta_2=\delta_3=\infty$). If CPLEX reports an infeasible solution, or has not generated an incumbent solution in 1800 seconds (30 minutes), we set δ_0 to the next value and try to solve again. Eventually, we will solve the model, or at least have an incumbent solution, so that we know that there is a feasible solution.
7. We now fix that δ_0 value and move onto the next division and repeat the process.
8. After carrying out this process for each division, we will have four δ_r values that we can use to solve the model.
9. A similar process is repeated for the four fixture case but the initial index into the distance vectors is element four, rather than element two.

We believe that this process has the benefit that as we consider the Premiership first, this will establish the lowest maximum distance for that division. This seems the right thing to do as more fans are affected by the Premierships teams (as they have larger fan bases, larger stadiums, attract more media interest etc.) so minimizing their distances first seems worthwhile.

The δ_r values we derived are shown in Table 14.

Table 14 Maximum distances for each division using the process presented in Section 5.5

	Prem		**Champ**		**Div 2**		**Div 3**	
Season	**Two** (δ_0)	**Four** (δ_0)	**Two** (δ_1)	**Four** (δ_1)	**Two** (δ_2)	**Four** (δ_2)	**Two** (δ_3)	**Four** (δ_3)
2002-2003	128	144	117	127	153	162	124	168
2003-2004	104	153	124	148	134	210	116	202
2004-2005	128	170	199	214	109	183	106	147
2005-2006	128	153	160	209	111	166	134	199
2006-2007	135	150	154	212	109	176	143	183
2007-2008	142	152	153	202	124	150	86	141
2008-2009	143	161	153	199	124	156	145	180

5.6 Results for Experiment 3: δ_r = maximum for each division

For the two day case (see Table 15), this experiment manages to produce similar solutions to the other experiments, in that solutions within 1% of optimality are obtained, with the exception of the 2002-2003 season. The four fixture case is more challenging (see Table 16). Only two seasons could generate a solution within 300 seconds. If we allow 7200 seconds, a solution was always returned and, similar to the other experiments, the solutions were about 15% of the lower bound. The solutions, with regard to the overall distance, are slightly higher than the other experiments but as we will discuss in the next section, this is not necessarily a problem.

Table 15 Results: Experiment 3a: δ_r = ***maximum for each division*** **(two day case:** $H = \{1,2\}$**)**

	300 Seconds			7200 Seconds			
Season	**LB**	**Found**	**% Gap**	**LB**	**Found**	**% Gap**	**Seconds**
2002-2003	4613.07	4965.09	7.09	4829.12	4958.09	2.62	2617
2003-2004	5339.89	5398.11	1.08	5360.32	5377.00	0.31	
2004-2005	5305.41	5345.10	0.74	5331.64	5345.10	0.25	
2005-2006	5081.13	5082.13	0.02	5081.13	5082.13	0.02	41
2006-2007	5325.86	5393.11	1.25	5365.08	5376.11	0.21	
2007-2008	5107.56	5153.12	0.88	5131.62	5132.12	0.01	362
2008-2009	5368.23	5385.10	0.31	5384.56	5385.10	0.01	1241

Table 16 Results: Experiment 3a: δ_r = ***maximum for each division*** **(four day case:** $H = \{1,2,3,4\}$**)**

	300 Seconds			7200 Seconds			
Season	**LB**	**Found**	**% Gap**	**LB**	**Found**	**% Gap**	**Seconds**
2002-2003	-	-	-	11525.04	14246.18	19.10	
2003-2004	-	-	-	12040.16	14465.22	16.76	
2004-2005	-	-	-	12110.68	14107.20	14.15	
2005-2006	12314.16	16909.26	27.18	12323.17	14284.26	13.73	
2006-2007	12367.57	16435.22	24.75	12511.40	14659.22	14.65	
2007-2008	-	-	-	12071.43	14471.24	16.58	
2008-2009	-	-	-	12525.00	14946.20	16.20	

6 Discussion

The results we reported in Section 5 are difficult to interpret, just by looking at the tables. In this section, we analyze the results for two seasons but they are representative of the underlying themes throughout the seven seasons (we summarize the other seasons in Section 6.3).

6.1 Season 2002-2003

We choose this season to analyze as it appears to be the most *difficult* season given that the gap is consistently over 1% whereas all other results (for the two day case) are under 1%. In Table 17 we present a summary of the various experiments. The table shows the total distance for the generated schedule, the maximum distance traveled by a team, the number of times that a team has to travel 180 miles or more (we chose this figure as 180 miles represents about three hours of driving time which seems a reasonable time limit for travel at this time of the year) and the number of Derby Clashes (i.e. when paired teams play each other). Our model actually treats Derby Clashes as a hard constraint (eq. 4), so for our experiments this value is always zero, but the published fixtures sometimes allow them.

Table 17 also shows the published fixture for the 2002-2003 season. The total distance was 7884 miles and the maximum distance for any one fixture was 171 miles (Newcastle vs Liverpool). No team had to travel over 180 miles but Rotherham and Sheffield Wednesday (which are paired) played each other. Having paired teams play each other is often beneficial, as far as minimizing the distance is concerned, as they are often local derbies and, by definition, the teams are close to each other (the distance between Rotherham and Sheffield Wednesday is 7 miles).

In all cases our model (we would suggest) is a significant improvement over the published fixtures (distances of under 5000 miles compared to 7884 miles). For our experiments, the maximum distance traveled by a single team is also an improvement over the published fixture (153 or 157 miles compared to 171 miles).

Choosing which experimental setup a user should choose would initially suggest 2a (as it has the lowest overall distance of 4862 miles) but we would urge caution. Experiment 3a sets a limit at the division level whereas both experiment 1a and 2a could allow greater distances, especially 1a, which allows infinite (of course the maximum distance is actually capped) travel distances. If we compare experiment 2a and 3a, we find that for the Premiership the total distance traveled is 908 miles (resp. 825 miles) for experiment 3a (resp. 2a). Therefore, it might appear that it would be more sensible to select experiment 2a as the methodology of choice. However, for the 3a experiment, as we set the maximum distance at the division level, no team had to travel more than 115 miles (Tottenham vs Aston Villa). For experiment 2a Southampton had to travel to Aston Villa (143 miles). Looking at the other divisions for experiment 2a, the maximum distances for each division are (we give all four divisions) Exp-2a=$\{143, 116, 153, 145\}$. The maximum distances for experiment Exp-3a=$\{115, 117, 153, 123\}$. For experiment 1a the values are Exp-1a=$\{143, 156, 157, 124\}$. Apart from a single mile (116 vs 117), experiment 3a produces the same, or lower, maximum distances than the other two experiments. Therefore, there would be a decision to be made. Does the problem owner want to minimize the total distance or take a more local view and ensure that no one club has to travel over a certain distance? There is no definitive answer as to which experiment returns the best result but as each experiment only takes five minutes, there is no reason why we cannot simply provide the user with all the solutions and let them decide which one is best.

Running the experiments for longer (7200 seconds) makes little difference to the overall results (see Table 18). The maximum distances for each division, for each experiment is as follows; Exp-1a={135, 117, 157, 124}, Exp-2a={143, 117, 153, 124}, Exp-3a={115, 117, 153, 124}. Exp-3a has the lowest (or equal) maximums across all divisions.

Table 17 Analysis: Season 2002-2003 (two day case, 300 seconds)

Experiment	Total Distance	Maximum Distance	# > 180	# of paired teams playing each other
Published	7884	171	0	1
1a (table 9)	4905	157	0	0
2a (table 12)	4862	153	0	0
3a (table 15)	4965	153	0	0

Table 18 Analysis: Season 2002-2003 (two day case, 7200 seconds)

Experiment	Total Distance	Max Distance	# > 180	# of paired teams playing each other
Published	7884	171	0	1
1a (table 9)	4801	157	0	0
2a (table 12)	4862	153	0	0
3a (table 15)	4958	153	0	0

For the four day case, we only consider the 7200 second experiment as we are likely to run the experiment for this amount of time if we were planning to use the results, as the 300 second experiment does not always return a solution. In fact, experiment 2b did not return a solution for the 7200 experiment so we cannot analyze it here.

The summary is presented in Table 19. We do not show the published fixtures as this season is classified as a two fixture season, so no data is available. Similar to the two days case, experiment 3b has a larger overall distance (14246 cf 13813) but has no fixtures that require a team to travel 180 miles or more. By comparison, experiment 1b has four fixture that requires teams to travel 180 miles or more. In fact, for experiment 3b, the maximum distance is only 168 miles. Experiment 1b has five fixtures greater than this, the four above 180 miles and another of 169 miles.

We note that the maximum distances for each division are as follows; Exp-1b={152, 166, 210, 202}; Exp-3b={144, 124, 162, 168}. Experiment 3b returns the lowest maximums across all four divisions.

Table 19 Analysis: Season 2002-2003 (four day case, 7200 seconds)

Experiment	Total Distance	Max Distance	# > 180	# of paired teams playing each other
1b (table 9)	13813	210	4	0
2b (table 12)	-	-	-	-
3b (table 15)	14246	168	0	0

6.2 Season 2005-2006

We chose to analyze the 2005-2006 season as this appears to be the *easiest* season as the gap in Table 9 is the lowest (0.01%) of all the seasons. However, it would appear that the football authorities had problems scheduling these fixtures as for the two day case there were 17 teams (see Table 20) that had to travel 180 miles or more and for the four day case (see Table 21 and also Appendix B) there were 37 teams that had to travel 180 miles or more.

At first sight, the fixtures that we have generated for the two fixture case seem to be a lot better than the published fixtures. However, we need to bear in mind that the 2005-2006 season is classified as a four fixture season (see section 3.4) so it is not really a fair comparison. However, we give our results to enable others to compare against our results. We also note that of the two figures available (see Table 32) we take the lowest as a comparison (i.e. of 10,626 and 11,333, we report 10,626 in this analysis).

All the experiments produced similar results with the maximum travel distance being either 160 or 161 miles. Experiment 3a produced a slighter higher overall distance (5082 miles) but, like 2002-2003 we are guaranteed to have a maximum distance traveled for each division. We have not shown the results for the 7200 second experiment, for the two day case, as the results are identical.

The maximum distances for each division are as follows; Exp-1a={161, 160, 112, 134}; Exp-2a={153, 160, 112, 134}; Exp-3a={128, 160, 111, 134}. Experiment 3a returns the lowest (or equal) maximums across all for divisions.

The more interesting analysis is for the four fixture schedule. These results are summarized in Table 21. We only provide the results for the 7200 second experiments as the gap tends to be quite large when we only allow 300 seconds, if a solution is found at all.

All the experiments give superior results to the published fixtures. Experiment 3b only has four fixtures where teams have to travel more 180 miles or more, whereas

Table 20 Analysis: Season 2005-2006 (Two day case, 300 seconds)

Experiment	Total Distance	Max Distance	# > 180	# of paired teams playing each other
Published	10626	304	17	3
1a (table 9)	5038	161	0	0
2a (table 12)	5040	160	0	0
3a (table 15)	5082	160	0	0

all the other solutions have at least eight teams traveling 180 miles or more and the published fixtures has 37 teams. This does come at the expense of a slightly higher overall total. If we look at the best solution, with regard to overall distance, (13785) the result from experiment 3b is 499 miles higher. This represents an average of just under three extra miles across the 184 fixtures. It would be up to the problem owner to make the final judgement which solution they prefer.

We note that the maximum distances for each division are as follows; Exp-1b={200, 220, 233, 219}; Exp-2b={203, 209, 182, 207}; Exp-3b={153, 209, 165, 199}. Experiment 3b returns the lowest (or equal) maximums across all for divisions.

Table 21 Analysis: Season 2005-2006 (Four day case, 7200 seconds)

Experiment	Total Distance	Max Distance	# > 180	# of paired teams playing each other
Published	21959	352	37	4
1b (table 10)	13785	233	8	0
2b (table 13)	13884	209	8	0
3b (table 16)	14284	209	4	0

6.3 Other Seasons

For completeness, we provide the summary tables, for the seasons not analyzed above. As above, if the season was not classified as a four season fixture we cannot provide the published fixture figures, although we still calculate our values for that season. Also, similar to above, when a season is classified as a four fixture season, we use the two fixture schedule that we indicated in the relevant section. We note again that this is not really a fair comparison with the published fixtures, but we are using the minimum distances in order to be as fair as possible.

6.3.1 2003-2004

As noted above, for the two day case, we are using an overall distance total of 8179 (see Section 3.2). As this season has been classified as a two fixture season, we cannot provide the statistics for the published fixture for the four day case, but we still generate our own set of fixtures.

Tables 22 and 23 summarizes the results for the 2003-2004 season.

The maximum distances for each division, for the two day case, are as follows; Exp-1a={188, 156, 134, 165}; Exp-2a=128, 128, 134, 134; Exp-3a={104, 124, 134, 116}. Experiment 3a returns the lowest (or equal) maximums across all four divisions

The maximum distances for each division, for the four day case, are as follows; Exp-1b={158, 157, 248, 243}; Exp-2b=170, 159, 210, 202; Exp-3b={153, 148, 210, 202}. Experiment 3b returns the lowest (or equal) maximums across all for divisions.

Table 22 **Analysis: Season 2003-2004 (Two day case, 7200 seconds)**

Experiment	Total Distance	Max Distance	# $>$ 180	# of paired teams playing each other
Published	8179	210	3	0
1a (table 9)	5209	188	1	0
2a (table 12)	5311	134	0	0
3a (table 15)	5377	134	0	0

Table 23 **Analysis: Season 2003-2004 (Four day case, 7200 seconds)**

Experiment	Total Distance	Max Distance	# $>$ 180	# of paired teams playing each other
1b (table 10)	13966	248	4	0
2b (table 13)	13805	210	3	0
3b (table 16)	14465	210	3	0

6.3.2 2004-2005

Tables 24 and 25 summarizes the results for the 2004-2005 season.

The maximum distances for each division, for the two day case, are as follows; Exp-1a={162, 219, 109, 124}; Exp-2a=161, 199, 109, 118; Exp-3a={128, 199, 209, 106}. Experiment 3a returns the lowest (or equal) maximums across all for divisions

The maximum distances for each division, for the four day case, are as follows; Exp-1b={200, 242, 208, 182}; Exp-2b=165, 214, 197, 144; Exp-3b={165, 214, 183, 147}. Experiment 3b returns the lowest (or equal) maximums across all for divisions

Table 24 Analysis: Season 2004-2005 (Two day case, 7200 seconds)

Experiment	Total Distance	Max Distance	# > 180	# of paired teams playing each other
Published	11012	257	9	0
1a (table 9)	5161	219	1	0
2a (table 12)	5212	199	1	0
3a (table 15)	5345	199	1	0

Table 25 Analysis: Season 2004-2005 (Four day case, 7200 seconds)

Experiment	Total Distance	Max Distance	# > 180	# of paired teams playing each other
Published	21959	???	??	?
1b (table 10)	13605	242	8	0
2b (table 13)	13782	214	8	0
3b (table 16)	14107	214	6	0

6.3.3 2006-2007

As noted above, for the two day case, we are using an overall distance total of 8439 (see Section 3.5). For the four day case, we are using a distance total of 23667

Tables 26 and 27 summarizes the results for the 2006-2007 season.

The maximum distances for each division, for the two day case, are as follows; Exp-1a={135, 157, 109, 150}; Exp-2a=150, 154, 137, 150; Exp-3a={135, 154, 109, 143}. Experiment 3a returns the lowest (or equal) maximums across all for divisions

The maximum distances for each division, for the four day case, are as follows; Exp-1b={174, 291, 202, 207}; Exp-2b=175, 212, 194, 199; Exp-3b={150, 212, 176, 183}. Experiment 3b returns the lowest (or equal) maximums across all for divisions

6.3.4 2007-2008

As noted above, for the two day case, we are using an overall distance total of 8644 (see Section 3.6). For the four day case, we are using a distance total of 22713

Tables 28 and 29 summarizes the results for the 2007-2008 season.

Table 26 Analysis: Season 2006-2007 (Two day case, 7200 seconds)

Experiment	Total Distance	Max Distance	# > 180	# of paired teams playing each other
Published	8439	213	4	1
1a (table 9)	5308	157	0	0
2a (table 12)	5358	154	0	0
3a (table 15)	5376	154	0	0

Table 27 Analysis: Season 2006-2007 (Four day case, 7200 seconds)

Experiment	Total Distance	Max Distance	# > 180	# of paired teams playing each other
Published	23667	364	42	2
1b (table 10)	14262	291	7	0
2b (table 13)	14619	212	8	0
3b (table 16)	14659	212	5	0

The maximum distances for each division, for the two day case, are as follows; Exp-1a={150, 154, 140, 86}; Exp-2a=148, 153, 140, 86; Exp-3a={142, 153, 124, 86}. Experiment 3a returns the lowest (or equal) maximums across all for divisions

The maximum distances for each division, for the four day case, are as follows; Exp-1b={205, 226, 176, 148}; Exp-2b=176, 202, 156, 141; Exp-3b={152, 202, 150, 141}. Experiment 3b returns the lowest (or equal) maximums across all for divisions

Table 28 Analysis: Season 2007-2008 (Two day case, 7200 seconds)

Experiment	Total Distance	Max Distance	# > 180	# of paired teams playing each other
Published	8644	213	1	1
1a (table 9)	5034	154	0	0
2a (table 12)	5096	153	0	0
3a (table 15)	5132	153	0	0

6.3.5 2008-2009

As noted above, for the two day case, we are using an overall distance total of 9312 (see Section 3.7). As this season has been classified as a two fixture season, we cannot provide the statistics for the published fixture for the four day case, but we still generate our own set of fixtures.

Tables 30 and 31 summarizes the results for the 2008-2009 season.

Table 29 Analysis: Season 2007-2008 (Four day case, 7200 seconds)

Experiment	Total Distance	Max Distance	# > 180	# of paired teams playing each other
Published	22713	311	17	2
1b (table 10)	14089	226	6	0
2b (table 13)	14011	202	2	0
3b (table 16)	14471	202	2	0

The maximum distances for each division, for the two day case, are as follows; Exp-1a={143, 190, 152, 199}; Exp-2a=143, 153, 152, 145; Exp-3a={143, 153, 124, 145}. Experiment 3a returns the lowest (or equal) maximums across all for divisions

The maximum distances for each division, for the four day case, are as follows; Exp-1b={192, 307, 208, 213}; Exp-2b=171, 190, 184, 188; Exp-3b={161, 199, 156, 180}. Experiment 3b returns the lowest (or equal) maximums for divisions, except for one, where it returns a maximum of 199 in experiment 3b, whereas experiment 2b returned a maximum of 190. This season's results are also different in that experiment 3b returns the lowest overall total.

Table 30 Analysis: Season 2008-2009 (Two day case, 7200 seconds)

Experiment	Total Distance	Max Distance	# > 180	# of paired teams playing each other
Published	9312	189	2	1
1a (table 9)	5244	199	2	0
2a (table 12)	5365	153	0	0
3a (table 15)	5385	153	0	0

Table 31 Analysis: Season 2008-2009 (Four day case, 7200 seconds)

Experiment	Total Distance	Max Distance	# > 180	# of paired teams playing each other
1b (table 10)	14671	307	7	0
2b (table 13)	14962	190	4	0
3b (table 16)	14946	199	4	0

7 Conclusion

We have presented a single model that minimizes the distance traveled by football supporters over the holiday season. The model is able to produce two or four complete fixtures, depending on the requirements of the football authorities. Several experiments were conducted, varying the parameters of the model. The model is able to produce solutions which are superior to those that are currently used. Previous discussions with the football league have suggested that we meet all the requirements, but it would be useful to hold further discussions with the authorities, as well as the police, in order to establish whether the model needs further refinement.

Of the three experiments that we conducted we would suggest that the option to limit each division to a maximum travel distance would probably be the most suitable to be used in practise as, although it usually produces slightly higher total distances, the solutions produced would probably be seen as being fairer when viewed by the supporters.

For our future work, the model presented in this chapter opens up the possibility to carry out more in-depth and *what-if* analysis on the fixtures for the holiday period. For example, are we able to reduce pair clashes whilst still minimizing the distance. If this is possible it could make significant savings for the police as they will not have to devote the same amount of resources to police the fixtures. We would also like to investigate weighting each pair clash, and including that in the objective function. This would be interesting as, at the moment, a pair clash between Liverpool and Everton, for example, is the same as a pair clash between Liverpool and Tranmere. However, the police would rather have Liverpool and Tranmere playing at home on the same day as this will be easier to police that the *Merseyside Derby*.

We will also turn our attention to generating schedules for the entire season. Given the experiences reported in this chapter, we do not believe that we will be able to produce optimal solutions and we feel that a (meta-)heuristic approach will be required.

Appendix

Summary of published fixtures and previous results

The 2005-2006 fixtures where supporters had to travel 180 miles or more

This appendix lists the 37 fixtures from the published fixture for the 2005-2006 season where a team is required to travel 180 miles or more. The figure in brackets is the distance in miles.

1. Darlington vs Torquay (352)

Table 32 This table summaries the distances traveled for the published fixtures (i.e. those that were actually played) and also the two fixture schedules that were generated in two previous papers (Kendall (2008); Kendall et al (2010b)) in order to provide a comparison with the results reported here. Note that there are slight differences from the figures shown in Kendall (2008) for 2002-2003 (7791 cf 7784), 2003-2004 (8168 cf 8179) and 2005-2006 (10631 cf 10626) due to minor errors found in the input data. Where N/A is specified, this indicates that this season did not produce a four fixture schedule. Only two fixture distances are shown for Kendall (2008); Kendall et al (2010b) as these papers did not generate four fixture schedules, with ** indicating that that paper did not generate schedules for those seasons.

	Published		**Kendall (2008)**	**Kendall et al (2010b)**
Season	**Two Day**	**Four Day**	**Two Day**	**Two Day**
2002-2003	7784	N/A	6040	5243
2003-2004	8179	N/A	6359	5464
2004-2005	11,012/12,261	23,273	6784	5365
2005-2006	10,626/11,333	21,959	6917	5234
2006-2007	11,761/11,906	23,667	**	5713
2007-2008	11,402/11,311	22,713	**	5366
2008-2009	9312	N/A	**	5564

2. Plymouth Argyle vs Leeds (321)
3. Plymouth Argyle vs Preston (304)
4. Newcastle vs Charlton (287)
5. Fulham vs Sunderland (281)
6. Tottenham vs Newcastle (280)
7. Hartlepool vs Southend (276)
8. Blackpool vs Southend (271)
9. Blackburn vs Portsmouth (268)
10. Hartlepool vs Swindon (267)
11. Bournemouth vs Scunthorpe (258)
12. Stockport vs Torquay (257)
13. Bournemouth vs Barnsley (242)
14. Norwich vs Preston (237)
15. Huddersfield vs Gillingham (235)
16. Swansea vs Gillingham (233)
17. QPR vs Burnley (232)
18. Darlington vs Barnet (230)
19. Norwich vs Burnley (229)
20. Tranmere vs Yeovil (229)
21. Doncaster vs Yeovil (228)
22. Everton vs Charlton (227)
23. Torquay vs Rushden & D'monds (224)
24. West Ham United vs Wigan (214)
25. Boston vs Carlisle (212)
26. Manchester City vs Chelsea (210)
27. Wolverhampton vs Plymouth Argyle (209)

28. Bradford vs Brentford (205)
29. Southampton vs Sheffield United (200)
30. Torquay vs Wycombe (199)
31. Manchester City vs Tottenham (199)
32. Arsenal vs Man Utd (197)
33. Stoke vs Ipswich (193)
34. Hull vs Ipswich (192)
35. Grimsby vs Carlisle (187)
36. Colchester vs Scunthorpe (185)
37. Brentford vs Swansea (182)

Number of allowed pairing violations

Table 33 shows the number of pairing violations that were present in the published fixtures. We allow ourselves the same number of violations in our solutions. Note that if (for example) Manchester United and Manchester are both playing at home, this counts as one pairing violation. In Kendall (2008) the same counts were used but each violation was counted as two. Further note that the number of pairing violations allowed for the four fixture schedule is simply double that of the two fixture schedule.

Table 33 Number of Allowed Pairing Violations.

Year	γ (Two fixtures)	γ (Four fixtures)
2002-2003	9	18
2003-2004	11	22
2004-2005	10	20
2005-2006	13	26
2006-2007	11	22
2007-2008	12	24
2008-2009	10	20

References

Anagnostopoulos A, Michel L, van Hentenryck P, Vergados Y (2006) A simulated annealing approach to the traveling tournament problem. Journal of Scheduling 9:177–193, DOI 10.1007/s10951-006-7187-8

Costa D (1995) An evolutionary tabu search algorithm and the nhl scheduling problem. INFOR 33:161–178

Crauwels H, van Oudheusden D (2002) A generate-and-test heuristic inspired by ant colony optimization for the traveling tournament problem. In: Burke EK, Causmaecker PD (eds) Proceedings of the 4th International Conference on the Practice and Theory of Automated Timetabling (PATAT 2002), pp 314–315

Dinitz J, Lamken E, Wallis W (2006) CRC Handbook of Combinatorial Designs, CRC Press, chap Scheduling a Tournament. (a previous edition of this book was published in 1995)

Easton K, Nemhauser G, Trick M (2001) The traveling tournament problem description and benchmarks. Principles and Practice of Constraint Programming - CP 2001, 7th International Conference, CP 2001, Paphos, Cyprus, November 26 - December 1, 2001, Lecture Notes in Computer Science 2239:580–584, iSBN: 978-3-540-42863-3 ISSN: 0302-9743 (Print) 1611-3349 (Online)

Easton K, Nemhauser G, Trick M (2003) Solving the traveling tournament problem: A combined integer programming and constraint programming approach. Lecture Notes in Computer Science 2740:100–109, DOI 10.1007/b11828, a previous version of this paper was published in the PATAT conference

Easton K, Nemhauser G, Trick M (2004) Sports Scheduling, Chapman & Hall, chap 52, pp 52–1 – 52–19

Elf M, Jnger, Rinaldi31 G (2003) Minimizing breaks by maximizing cuts. Operations Research Letters 31:343–349

Kendall G (2008) Scheduling English football fixtures over holiday periods. Journal of the Operational Research Society 59(6):743–755

Kendall G, Knust S, Ribeiro C, Urrutia S (2010a) Scheduling in sports: An annotated bibliography. Computers & Operations Research 37:1–19

Kendall G, McCollum B, Cruz F, McMullan P (2010b) Scheduling English football fixtures: Consideration of two conflicting objectives. In: McCollum B, Burke E (eds) Proceedings of the 8th International Conference on the Practice and Theory of Automated Timetabling (PATAT 2010), pp 1–15

Knust S (2009) Classification of literature on sports scheduling. Available online at http://www.inf.uos.de/knust/sportssched/sportlit_class/, last accessed on October 15th 2009

Rasmussen R, Trick MA (2008) Round robin scheduling a survey. European Journal of Operational Research 188:617–636

Ribeiro C, Urrutia S (2004) Heuristics for the mirrored traveling tournament problem. In: Proceedings of the 5th International Conference on the Practice and Theory of Automated Timetabling (PATAT 04), pp 323–342

Rollin R, Rollin J (eds) (2002) Rothmans Football Yearbook 2002–2003. Headline Book Publishing: London, UK, iSBN: 0-7553-1100-0

Rollin R, Rollin J (eds) (2003) Sky Sports Football Yearbook 2003–2004. Headline Book Publishing: London, UK, iSBN: 0-7553-1228-7

Rollin R, Rollin J (eds) (2004) Sky Sports Football Yearbook 2004–2005. Headline Book Publishing: London, UK, iSBN: 0-7553-1311-9

Rollin R, Rollin J (eds) (2005) Sky Sports Football Yearbook 2005–2006. Headline Book Publishing: London, UK, iSBN: 0-7553-1385-2

Rollin R, Rollin J (eds) (2006) Sky Sports Football Yearbook 2006–2007. Headline Book Publishing: London, UK, iSBN: 0-7553-1526-X

Rollin R, Rollin J (eds) (2007) Sky Sports Football Yearbook 2007–2008. Headline Book Publishing: London, UK, iSBN: 0-7553-1664-9

Rollin R, Rollin J (eds) (2008) Sky Sports Football Yearbook 2008–2009. Headline Book Publishing: London, UK, iSBN: 0-7553-1820-9

Trick M (2009) Traveling tournament instances. Available online at http://mat.gsia.cmu.edu/TOURN. last accessed on October 16th 2009

Urrutia S, Ribeiro C (2004) Minimizing travels by maximizing breaks in round robin tournament schedules. Electronic Notes in Discrete Mathematics 18-C:227–233

de Werra D (1981) Studies on Graphs and Discrete Programming, North-Holland, chap Scheduling in sports, pp 381–395

de Werra D (1988) Some models of graphs for scheduling sports competitions. Discrete Applied Mathematics 21:47–65

Westphal S (2011) Scheduling the german basketball league. Under review

Westphal S, Noparlik K (2010) A 5.875-approximation for the traveling tournament problem. In: Proceedings of the Practice and Theory of Automated Timetabling (PATAT 2010), pp 417–426