American Review of Mathematics and Statistics
December 2015, Vol. 3, No. 2, pp. 7-14
ISSN: 2374-2348 (Print), 2374-2356 (Online)

Published by American Research Institute for Policy Development
DOI: 10.15640/arms.v3n2a2
URL: http://dx.doi.org/10.15640/arms.v3n2a2

A Mathematical Model for the Traveling Tournament Problem

J. Mark Dorrepaal[1], George Chackman[2] & Brian Dorrepaal[3]

Abstract

The traveling tournament problem (TTP) is a sports scheduling problem which has generated considerable interest over the last decade. In this paper a mathematical model is introduced in which the home venues of all teams lie at equal intervals along a straight line. This model is applied to the six team problem and brute force methods are used to find all optimal schedules. These schedules are analyzed and categorized into equivalence classes. A number of schedules which have interesting properties are printed in full.

Keyword: Sports scheduling, traveling tournament problem, double round robin tournament, Minimum distance, and combinatorial optimization problem

1. Introduction

The Traveling Tournament Problem (TTP) is a sports scheduling problem requiring construction of a minimum distance double round robin tournament for a group of n teams where n is even and greater than 2. The problem was first defined in a pair of papers by Easton, Nemhauser and Trick (2001, 2003). The motivation for studying TTP originated with Major League Baseball (MLB) which was interested in improving the travel efficiency of their team schedules. Trick (2002) established a website with a challenge page which solicited solutions to different versions of the problem. The original problem defined at the website involved TTPs for $n = 4, 6, 8, \ldots, 16$ where the home venues for the n teams were cities in MLB's National League. These TTPs are called the NLn instances.

Since its introduction in 2001, much of the interest in the TTP has focused on developing and refining numerical schemes to attack the problem. These have produced steady improvements in schedule efficiency. The biggest advances along these lines have come from Anagnostopoulos, Michel, Van Hentenryck and Vergados (2006) who used a simulated annealing algorithm to obtain many superior results in the larger NLn instances.

In this paper we will study only the 6-team TTP, but we will introduce a simple geometric model for the placement of the 6 home venues. This model will greatly facilitate our analysis and understanding of optimal solutions to the 6-team problem. It will enable the use of brute force numerical methods to quantify the total number of legal 6-team schedules and to discover properties shared by all optimal schedules.

2. The Constrained Traveling Tournament Problem

The TTP considered in this paper is subject to the same constraints imposed on the NL6 instances at Trick's website (2002). We seek a double round robin tournament (DRRT) of 6 teams in which each team plays exactly once in each of 10 slots. No team can play more than three consecutive home or away games (At-most constraint). A team cannot play the same opponent in consecutive slots (No-repeater constraint). The objective is to minimize the total distance traveled by the 6 teams collectively.

[1] Dept. of Mathematics & Statistics, Old Dominion University, Norfolk, VA 23529, Tel.: +1-757-6833908, Fax: +1-757-6833885, E-mail: jdorrepa@odu.edu

[2] 2032E Strome Entrepreneurial Center, Old Dominion University, Norfolk, VA 23529, E-mail: gchackma@odu.edu

[3] E-mail: bdorre@sympatico.ca

We propose a geometric model in which the 6 teams are named A, B, C, D, E and F and are placed at equal intervals along a straight line. In particular, we take team A's home venue to be located at $x = 1$ on the number line, team B's home venue is at $x = 2$, team C's home is at $x = 3$, etc. We refer to these distances as "miles". Therefore each team is exactly one "mile" from its nearest neighbor(s).

The distance matrix M_6 is a symmetric 6 × 6 matrix which gives the distance between the team occupying the i^{th}row and the team occupying thej^{th} column. The distance matrix M_6 is given in Table 1.

Table 1: The distance matrix for a 6-team TTP

	A	B	C	D	E	F
A	0	1	2	3	4	5
B	1	0	1	2	3	4
C	2	1	0	1	2	3
D	3	2	1	0	1	2
E	4	3	2	1	0	1
F	5	4	3	2	1	0

3. The Independent Lower Bound

In their 2003 paper, Easton et al. define a team's optimal tour: "A team's optimal tour minimizes the travel distance for that team exclusive of the other teams in the schedule. The sum over n teams of the distances associated with their optimal tours provides a simple but strong lower bound on the TTP. We call this the Independent Lower Bound (ILB)".

In the current model the ILB is simple to calculate. Because of the at-most constraint, all teams must make two road trips to complete their away-game schedule obligations. Team A is a geographically remote team and should take a three-game road trip to D, E, F, and a second two-game trip to B, C. The three-game trip covers a round-trip distance of 10 miles (provided D is <u>not</u> the second stop on the trip), and the two-game trip is 4 miles. Therefore A's optimal tour is 14 miles. In a similar way it can be shown that B's optimal tour is 12 miles and C's is 10 miles. The geometric symmetry built into the model leads to similar optimal tours for teams F, E, D respectively. The ILB for our 6-team TTP is $14 + 12 + 10 + 10 + 12 + 14 = 72$ miles.

4. Schedule Format

Schedules can be presented in an alphabetic format or an equivalent numerical format. Both schedule types involve a 10 × 6 matrix whose columns correspond to the 6 teams A, B, C, D, E, F and whose rows give the "days" or slots over which the tournament is held. In the alphabetic format, the team letters are used with road games indicated by the symbol "@". In the numerical format, the teams are referred to by number (A=1, B=2, C=3, etc.) and road games are indicated by a minus sign. Tables 2 and 3 give the same schedule in the two formats.

Table 2: A six team schedule in alphabetic format

Row	A	B	C	D	E	F
1	B	@A	@D	C	F	@E
2	E	C	@B	F	@A	@D
3	@F	E	D	@C	@B	A
4	@E	D	@F	@B	A	C
5	D	@C	B	@A	@F	E
6	C	F	@A	E	@D	@B
7	F	@D	E	B	@C	@A
8	@D	@E	F	A	B	@C
9	@C	@F	A	@E	D	B
10	@B	A	@E	@F	C	D

Table 3: The same schedule in numerical format

1	2	3	4	5	6
2	-1	-4	3	6	-5
5	3	-2	6	-1	-4
-6	5	4	-3	-2	1
-5	4	-6	-2	1	3
4	-3	2	-1	-6	5
3	6	-1	5	-4	-2
6	-4	5	2	-3	-1
-4	-5	6	1	2	-3
-3	-6	1	-5	4	2
-2	1	-5	-6	3	4

5. Search for Optimal Schedules

The schedule in Table 2 is made up of 10 slots. There are a total of 120 different slot choices from which these 10 can be chosen. As we begin to assemble a six team schedule, the first row can be chosen in 120 different ways. Because of the no-repeater constraint, there are only 64 slot choices for row 2. For row 3, there are only 45 slot choices. This means that for all 6-team legal schedules, the first three rows of the schedule can occur in $120 \times 64 \times 45 = 345{,}600$ different ways. When we move to the fourth row of the schedule, the analysis becomes more complicated because the at-most constraint now becomes a factor. If any of the six teams have played all of their first three games at home or on the road, then the at-most constraint will dictate a change in venue for such teams. This will have a limiting effect on the number of slot choices for row 4 of the schedule. It can be shown, however, that the maximum number of slot choices for row 4 will be 30 and the maximum number for row 5 will be 18.

These numbers suggest that the total number of legal 6-team schedules is at least of order 10^8 and probably in the billions. This is not so large that it cannot be attacked by brute force methods. A code was developed which constructed legal 6-team schedules row by row in a fashion which guaranteed the enforcement of both no-repeater and at-most constraints. Whenever a full 10-row schedule was completed, it was counted and its distance calculated. It soon became apparent that an optimal 6-team schedule had a distance of 84 miles – twelve miles over the Independent Lower Bound. Armed with this statistic, the program was modified to count all legal schedules and to save only the optimal ones. A successful run was made which took 100 hours of computing time.

6. Results

There are exactly $6{,}531{,}327{,}360$ legal 6-team TTP schedules. This result is independent of the model we are using and therefore we can conclude that this value defines the total number of NL6 instances as well. The numerical work also produced 494 optimal schedules. For every optimal schedule S, there was also found the upside-down schedule S^{-1} which consists of exactly the same rows of schedule S, but in reverse order. Clearly S^{-1} is also optimal. If we identify these two schedules as equivalent, then we have 247 independent schedules. An interesting feature of these optimal schedules is that in every case, team A either begins or ends its schedule by playing its closest rival B. This is also true of team F. In every optimal schedule, team F either begins or ends its schedule by playing team E. Focusing our attention on team A, we can decide which schedule, S or S^{-1}, we will use to categorize the set of all optimal schedules. For each pair of inverse schedules, we will choose S to be the schedule in which team A plays team B in the first game (row 1). For schedules in which A plays B in rows 1 and 10, we define S to be the schedule where A's first game is B@A with A being the home team. Using this convention we find that of the 247 optimal schedules, 82 of them have A playing B in both rows 1 and 10. Clearly the first game is B@A and the last game is A@B.

Furthermore, there are 123 schedules where A's first game is B@A and A's last game is against an opponent other than B. Finally, there are 42 schedules where A's first game is A@B and A's last game is against either C or D or E or F. These results are summarized in Table 4.

Table 4: Categorization of Optimal Schedules

Category	# of Schedules	Approx. Fraction of Total = 247
(+2, –2)	82	1/3
(+2, *x*)	123	1/2
(–2, *x*)	42	1/6

In Table 4, the category is an ordered pair whose first component is team A's first opponent using the numerical format introduced in Table 4. The second component of the ordered pair is team A's final opponent. Thus category (+2, –2) includes all optimal schedules where A (team 1) hosts B (team 2) in its first game, and visits B (team 2) in its final game. Category (+2,*x*) includes all optimal schedules where A hosts B in its first game and plays a different team (either home or away) in its final game. Similarly for category (–2, *x*). Table 4 indicates that roughly $\frac{1}{2}$ of the optimal schedules belong to category (+2, *x*), roughly $\frac{1}{3}$ are in (+2,–2) and the remaining $\frac{1}{6}$ are (–2, *x*).

7. Equivalence Classes of Optimal Schedules

In addition to the distinction that can be made between a schedule S and its inverse S^{-1}, there is another grouping that can be discussed. Given an optimal schedule S, the permutation mapping (A,B,C,D,E,F) → (F,E,D,C,B,A) will transform S into another schedule S^* which is also optimal because of the geometric symmetry of the home venues of the six teams. We will call S^* the reflected version of S. The schedule in Table 2 is an optimal schedule in the category (+2,–2). Its reflected schedule S^* is given in Table 5 (alphabetic format) and in Table 6 (numerical format). Table 5 also includes the distances traveled by each of the 6 teams, the optimal tours for each team and the differences between these values.

Table 5: The reflected version of the schedule in Table 2

Row	A	B	C	D	E	F
1	@B	A	D	@C	@F	E
2	@C	@F	A	@E	D	B
3	F	@E	@D	C	B	@A
4	D	F	@E	@A	C	@B
5	B	@A	@F	E	@D	C
6	@E	@C	B	@F	A	D
7	@F	@D	E	B	@C	A
8	@D	E	F	A	@B	@C
9	E	C	@B	F	@A	@D
10	C	D	@A	@B	F	@E
Distance	14	14	10	18	12	16
Optimal Tour	14	12	10	10	12	14
Difference	0	2	0	8	0	2

Table 6: The reflected version of the schedule in Table 3

1	2	3	4	5	6
-2	1	4	-3	-6	5
-3	-6	1	-5	4	2
6	-5	-4	3	2	-1
4	6	-5	-1	3	-2
2	-1	-6	5	-4	3
-5	-3	2	-6	1	4
-6	-4	5	2	-3	1
-4	5	6	1	-2	-3
5	3	-2	6	-1	-4
3	4	-1	-2	6	-5

If $S(k,n)$ is the 10×6 matrix representing schedule S in numerical format, $S^{-1}(k,n)$ is the corresponding inverse schedule and $S^*(k,n)$ is the reflected schedule, then

$S^{-1}(k,n) = S(11-k,n)$ (1)
$S^{*}(k,n) = 7\frac{S(k,7-n)}{|S(k,7-n)|} - S(k,7-n)$ (2)

where $1 \leq k \leq 10, 1 \leq n \leq 6$.

Given an optimal schedule S, we define the equivalence class of S to be the set of four schedules $Q = \{S, S^{-1}, S^{*}, (S^{*})^{-1}\}$. Clearly the operations of inversion and reflection preserve distance. Therefore all four members of Q are optimal. Equivalence classes provide a more economical way of categorizing optimal schedules.

8. Symmetric Schedules

Clearly there is a discrepancy between the fact reported in section 6 that there are 494 optimal schedules with the findings in section 7 that optimal schedules can be grouped into equivalence classes consisting of 4 schedules. The resolution of the problem lies in the fact that some optimal schedules are symmetric. A symmetric schedule is one for which $S^{*} = S^{-1}$; that is, its reflection is exactly the same as its inverse. Furthermore, if S is symmetric, then

$(S^{*})^{-1} = (S^{-1})^{*} = S$ (3)

An example of a symmetric schedule is given in Table 7 (alphabetic format) and in Table 8 (numerical format).
Let $S(k,n)$ be the 10×6 matrix representing a symmetric schedule in numerical format. Equating (1) and (2) yields the following criterion for a symmetric schedule

$S(k,n) + S(11-k,7-n) = 7\frac{S(k,n)}{|S(k,n)|}$ (4)

Table 7: A symmetric schedule of category (+2,–2)

Row	A	B	C	D	E	F
1	B	@A	@D	C	F	@E
2	F	@C	B	@E	D	@A
3	@D	F	E	A	@C	@B
4	@C	E	A	@F	@B	D
5	E	D	@F	@B	@A	C
6	D	@F	@E	@A	C	B
7	C	@E	@A	F	B	@D
8	@E	@D	F	B	A	@C
9	@F	C	@B	E	@D	A
10	@B	A	D	@C	@F	E
Distance	16	12	14	14	12	16
Optimal Tour	14	12	10	10	12	14
Difference	2	0	4	4	0	2

Table 8: The same schedule in numerical format

1	2	3	4	5	6
2	-1	-4	3	6	-5
6	-3	2	-5	4	-1
-4	6	5	1	-3	-2
-3	5	1	-6	-2	4
5	4	-6	-2	-1	3
4	-6	-5	-1	3	2
3	-5	-1	6	2	-4
-5	-4	6	2	1	-3
-6	3	-2	5	-4	1
-2	1	4	-3	-6	5

If a schedule S is symmetric, then its equivalence class Q_{sym} will contain only two schedules. We have $Q_{sym} = \{S, S^{-1}\}$ because $S^* = S^{-1}$ which is already included, and $(S^*)^{-1} = S$. It was found that of the 494 optimal schedules, 26 were symmetric. The remaining 468 were non-symmetric. Therefore there are 13 equivalence classes each possessing 2 symmetric schedules and 117 equivalence classes with 4 schedules each. In this way the set of all optimal schedules divides into 130 equivalence classes of which exactly 10% consist of symmetric schedules.

9. Unique Pairing Schedules

Another characteristic exhibited by the schedule in Table 7 is that it is a unique pairing (UP) schedule. Note that the rivalries in row 1 have B@A, C@D and F@E. Teams A and B do not play each other again until row 10 – but when they do, the rivalries are once again A@B, D@C and E@F. Thus in rows 1 and 10, the schedule exhibits the same rivalries but opposite venues. The same is true for rows 2 and 9, rows 3 and 6, rows 4 and 7, and for rows 5 and 8. In a UP schedule, for every row in the schedule matrix, there is a complementary row consisting of the same rivalries at opposite venues.

A non-unique pairing (XUP) schedule is illustrated in Table 2. Note that the rivalries in row 1 are the same as the schedule discussed above: B@A, C@D and F@E. In row 10, team A plays at B; but now team C plays @ E and D plays @ F. So the rivalries are different in rows 1 and 10 even though A and B play each other. A similar thing happens between rows 2 and 4. Teams A and E play each other, but the other four teams have different opponents in those two rows. The one exception to this pattern occurs in rows 6 and 9. In those two rows, the rivalries are the same and the venues are opposite. A XUP schedule is one where it is generally true that in the two rows where team A plays the same opponent, the remaining four teams play different opponents. All optimal XUP schedules have exactly two complementary rows (same rivalries and opposite venues) and no complementary pairs in the remaining eight rows.

Not all symmetric schedules are UP. We will examine some very interesting XUP symmetric schedules in section 11. Non-symmetric schedules also come in UP as well as XUP forms.

10. Categorizing Equivalence Classes

It was shown in section 8 that the optimal schedules fit into 130 equivalence classes. In section 6 we found that there are three categories of optimal schedules based on team A's opponent in its first and last games. We present here tables showing how these categories break down into equivalence classes and schedule type (UP or XUP). First consider the symmetric schedules. It was observed in section 8 that there are 13 symmetric equivalence classes consisting of two optimal schedules each. Of these, six are in the category (+2, –2), six are in (+2,*x*) and one is in (–2,*x*). Table 9 provides further analysis. In particular, all symmetric optimal schedules in category (+2, –2) are UP while those in the remaining two categories are XUP.

Table 9: Symmetric Equivalence Classes

Category	UP	XUP	TOT
(+2, –2)	6	0	6
(+2, *x*)	0	6	6
(–2, *x*)	0	1	1
	6	7	13

In order to extend our analysis to non-symmetric equivalence classes (each containing four optimal schedules), we must refine our categorization conventions. Given an optimal non-symmetric schedule S, its corresponding equivalence class will be $Q = \{S, S^{-1}, S^*, (S^*)^{-1}\}$. We define Q to belong to category (+2, –2) if <u>any</u> schedule within Q has team A playing its first and last games against team B. This definition is necessary because it can happen that schedule S is in category (+2,*x*), but schedule S^* is in (+2, –2). Furthermore once all members of category (+2, –2) have been identified, we define Q to belong to category (+2, x) if <u>any</u> schedule within Q has B@A as its first game. This definition is necessary because it can happen that schedule S is in category (–2,*x*), but schedule S^* is in (+2,*x*). In this way all non-symmetric equivalence classes can be categorized. The results are shown in Table 10.

Table 10: Non-Symmetric Equivalence Classes

Category	UP	XUP	TOT
(+2, –2)	28	20	48
(+2, *x*)	29	38	67
(–2, *x*)	0	2	2
	57	60	117

Every equivalence class in Table 9 contains two optimal schedules and every equivalence class in Table 10 contains 4. Therefore of the original 494 optimal schedules mentioned in section 6, exactly $2(6) + 4(57) = 240$ are UP and 254 are XUP. This is interesting because when considering all legal 6-team schedules, there are approximately 5.5 times as many XUP schedules as there are UP schedules. Clearly the order which is present in a UP schedule is more conducive to optimality. This is because the no-repeater constraint is guaranteed for all teams in a UP schedule provided team A satisfies the no-repeater constraint. In a XUP schedule, enforcement of the no-repeater constraint must be verified individually for all 6 teams.

11. Three Interesting Symmetric Schedules

In this section we present three optimal symmetric schedules with interesting distance configurations. As reported in section 5, the six teams in an optimal schedule travel a total of 84 miles. The schedule in Table 11 is one where every team travels the same distance – 14 miles. This schedule is a UP schedule in the category (+2,–2).

Table 11: All teams travel the same distance

Row	A	B	C	D	E	F
1	B	@A	D	@C	F	@E
2	C	@E	@A	F	B	@D
3	@D	@F	E	A	@C	B
4	@F	C	@B	E	@D	A
5	@E	D	@F	@B	A	C
6	D	F	@E	@A	C	@B
7	F	@C	B	@E	D	@A
8	E	@D	F	B	@A	@C
9	@C	E	A	@F	@B	D
10	@B	A	@D	C	@F	E
Distance	14	14	14	14	14	14
Optimal Tour	14	12	10	10	12	14
Difference	0	2	4	4	2	0

Table 12 exhibits a schedule where each team's distance exceeds its optimal tour by the same amount – 2 miles. Generally speaking the geographically remote teams like A and F will have to travel further distances to fulfil their schedule obligations. On top of this, all optimal schedules require the six teams collectively to exceed their optimal tours by 12 miles. The schedule in Table 12 spreads this net excess evenly amongst the six teams. The schedule is XUP and of category (–2,*x*). From Table 9 it is the only optimal symmetric schedule having these properties. Of further interest is the fact that team B's 3-game road trip is inefficient. This means that D is the second stop on this trip which makes the distance of this trip 10 miles rather than the usual 8 miles. An attempt to make this trip efficient by interchanging rows 2 and 3 leads to a violation of the at-most constraint for team E.

Table 12: Each team's distance exceeds optimal tour by 2 miles

Row	A	B	C	D	E	F
1	@B	A	@F	E	@D	C
2	@D	@F	@E	A	C	B
3	@C	@D	A	B	@F	E
4	F	@E	D	@C	B	@A
5	C	F	@A	@E	D	@B
6	@E	C	@B	@F	A	D
7	@F	E	@D	C	@B	A
8	B	@A	E	F	@C	@D
9	E	D	F	@B	@A	@C
10	D	@C	B	@A	F	@E
Distance	16	14	12	12	14	16
Optimal Tour	14	12	10	10	12	14
Difference	2	2	2	2	2	2

The schedule in Table 13 has the property that the majority of teams (4 of 6) travel their optimal tours. This symmetric schedule is XUP and is of category (+2, *x*). Of all the optimal schedules, this one (and its inverse) are the only ones where over half the teams travel their optimal tours.

Table 13: Four of six teams travel their optimal tours

Row	A	B	C	D	E	F
1	B	@A	@F	E	@D	C
2	C	D	@A	@B	F	@E
3	@E	C	@B	F	A	@D
4	@F	@E	D	@C	B	A
5	@D	@F	E	A	@C	B
6	E	@D	F	B	@A	@C
7	F	E	@D	C	@B	@A
8	@C	F	A	@E	D	@B
9	@B	A	@E	@F	C	D
10	D	@C	B	@A	@F	E
Distance	14	12	16	16	12	14
Optimal Tour	14	12	10	10	12	14
Difference	0	0	6	6	0	0

References

Anagnostopoulos, A., Michel, L., Van Hentenryck, P., Vergados, Y. (2006) A simulated annealing approach to the traveling tournament problem. J Sched 9, 177-193.

Easton, K., Nemhauser, G., Trick, M. (2001) The traveling tournament problem description and benchmarks. Lecture Notes in Computer Science, Vol. 2239 (580-584) DOI: 10.1007/3-540-45578-7_43.

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, Vol. 2740 (100-109) DOI: 10.1007/978-3-540-45157-0_6.

Trick, M. (2002) Challenge traveling tournament instances. http://mat.gsia.cmu.edu/TOURN/