

SPORTS SCHEDULING: PROBLEMS AND APPLICATIONS

CELSO C. RIBEIRO

ABSTRACT. Sports scheduling problems mainly consist in determining the date and the venue in which each game of a tournament will be played. Integer programming, constraint programming, metaheuristics, and hybrid methods have been successfully applied to the solution of different variants of this problem. This paper provides an introductory review of fundamental problems in sports scheduling and their formulations, followed by a survey of applications of optimization methods to scheduling problems in professional leagues of different sport disciplines such as football, baseball, basketball, cricket, and hockey. A case study illustrates a real-life application of integer programming to the schedule of the yearly Brazilian football tournament.

1. INTRODUCTION

Sports have become a big business in a global economy. Tournaments are followed by millions of people across the world. Teams make big investments in new players. Broadcast rights amount to hundreds of millions of dollars in some competitions. Countries and cities fight for the right to organize worldwide events such as the Olympics and the Football World Cup.

Professional sport leagues involve millions of fans and significant investments in players, broadcast rights, merchandising, and advertising, facing challenging optimization problems. On the other side, amateur leagues involve less investments, but also require coordination and logistical efforts due to the large number of tournaments and competitors.

The main problem in sports scheduling consists in determining the date and the venue in which each game of a tournament will be played. Applications are found in the scheduling of tournaments of sports such as football, baseball, basketball, cricket, and hockey. These problems have been solved by different exact and approximate approaches, including integer programming, constraint programming, metaheuristics, and hybrid methods.

There are many relevant aspects to be considered in the determination of the best schedule for a tournament. In some situations, one seeks for a schedule minimizing the total traveled distance, as in the case of the traveling tournament problem (Easton et al., 2001) and in that of its mirrored variant (Ribeiro and Urrutia, 2007b), which is common to many tournaments in South America (Durán et al., 2007b). Other problems attempt to minimize the total number of breaks, i.e., the number of pairs of consecutive home games or consecutive away games played by the same team. The minimization of the carry-over effects value (Russell, 1980) is another fairness criterion leading to an even distribution of the sequence of games

Date: Submitted January 2, 2011; Revised April 4, 2011; Accepted May 22, 2011.

Key words and phrases. Sports scheduling, tournaments, round robin, traveling tournament problem, referee assignment, heuristics, integer programming.

along the schedule. Some problems in sports scheduling have a multicriteria nature. Ribeiro and Urrutia (2007a; 2009) tackled the scheduling of the yearly Brazilian football tournament, preliminarily formulated as a bicriteria optimization problem in which one of the objectives consisted in maximizing the number of games that could be broadcast by open TV channels (to increase the revenues from broadcast rights) and the other consisted in finding a balanced schedule with a minimum number of home breaks and away breaks (for sake of fairness). A multicriteria version of a referee assignment problem arising in amateur leagues (Duarte et al., 2007a;b) was tackled by Duarte and Ribeiro (2008).

This paper provides an introductory review to the main problems in sports scheduling, also covering the principal practical applications. Although being more focused into problems and applications, it also addresses the main solution methods and innovative algorithmic approaches applied in their solution. It should be considered as a starting point for newcomers and research in the area. The interested reader is referred to Rasmussen and Trick (2008) for a comprehensive survey of the literature on round robin tournament scheduling and to Kendall et al. (2010) for a rather complete bibliography of scheduling problems in sports.

The remaining of this paper is organized as follows. Section 2 reviews the main definitions and basic issues. Section 3 presents an overview of the main problems arising in sports scheduling and their formulations: breaks minimization, distance minimization and the traveling tournament problem, and carry-over effects minimization. Problem reformulation techniques are investigated in Section 4, which explores a variant of the traveling tournament problem where the venues are known beforehand. Section 5 surveys applications in different sport disciplines like football, baseball, basketball, cricket, and hockey. A case study of a recent application of integer programming in the scheduling of the yearly first division football tournament in Brazil is reported in Section 6. Concluding remarks and references to other scheduling problems in sports are given in the last section.

2. DEFINITIONS

We consider a tournament played by an even number n of teams. A round robin tournament is one in which each team plays against every other a fixed number of times. Every team faces each other exactly once (resp. twice) in a single (resp. double) round robin (SRR) (resp. DRR) tournament and plays at most once in each round. A round robin tournament is said to be compact if the number of rounds is minimum and every team plays exactly once in every round. Each team has its own venue at its home city and each game is played at the venue of either one of the two teams in confrontation. The team that plays at its own venue is called the home team and is said to play a home game, while the other is called the away team and is said to play an away game. We say there is a repeater whenever the same pair of teams face each other twice in two consecutive rounds. If the number of teams is odd, then in each round one team has a bye, i.e. it does not play. This situation may be reduced to the case of an even number of teams by adding a dummy team. Then, in each round the team playing against the dummy team has a bye.

Double round robin tournaments are often partitioned into two phases, where each game has to occur exactly once in each phase, but with different home rights. In the case of the so-called mirrored schedules, the games played by each team in

the second phase follow exactly the same order as those played in the first, but with exchanged venues. Therefore, the two games played by each pair of opponents take place at the same round of the first and second phases.

Tournaments may be represented by graphs, which offer a good model for scheduling formulations and algorithms, see e.g. de Werra (1980; 1981; 1988). The complete graph K_n may be used to represent a single round robin tournament or any of the phases of a compact double round robin tournament. Each of its nodes represents a team. Each game is represented by an edge, whose extremities are associated with the two opponent teams. Figure 1 displays an example illustrating the graph representation of a single round robin tournament with $n = 4$ teams.

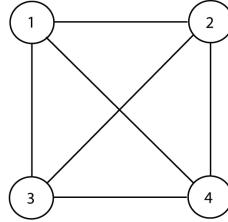


FIGURE 1. Example of a single round robin tournament with $n = 4$ teams represented by a complete graph.

An edge coloring with exactly $n - 1$ colors corresponds to an 1-factorization of K_n , i.e. a partitioning of its edge set into 1-factors F_1, \dots, F_{n-1} (each consisting of $n/2$ non-adjacent edges). Each 1-factor corresponds to the games scheduled to a given round. Therefore, an ordered 1-factorization determines a timetable for the tournament, defining the round in which each game is played. Figures 2(a) to 2(c) represent a timetable for the example presented in Figure 1, in which each game is assigned to a round.

A tournament schedule determines not only the round in which each game is played, but also its venue. If home and away games have to be distinguished, then an orientation is assigned to the edges of the complete graph and to its 1-factors. In the case of the example depicted in Figures 1 and 2, we assume that team 2 plays all its games away, team 1 plays away with team 3, team 4 plays away with team 1, and team 4 plays at home with team 3. Therefore, the four teams have the following home-away patterns of game playing: team 1 - home/away/home; team 2 - away/away/away; team 3 - away/home/home; and team 4 - home/home/away. Figure 3 illustrates this situation, in which an orientation has been assigned to each edge in Figure 1 (or, equivalently, to each edge in Figures 2(a) to 2(c)) to create a complete schedule defined by an oriented graph: the existence of an arc from node i to node j means that team i plays away against team j .

The problem of scheduling a round robin tournament is often divided into two subproblems. The construction of the timetable consists in determining the round in which each game will be played. The home-away pattern (HAP) set determines in which condition (home or away) each team plays in each round. The HAP set for the previous example can be represented by a matrix as in Table 1, in which the cell corresponding to row k and column j indicates the playing condition of team j in round k . Together, the timetable and the home-away pattern set determine the tournament schedule.

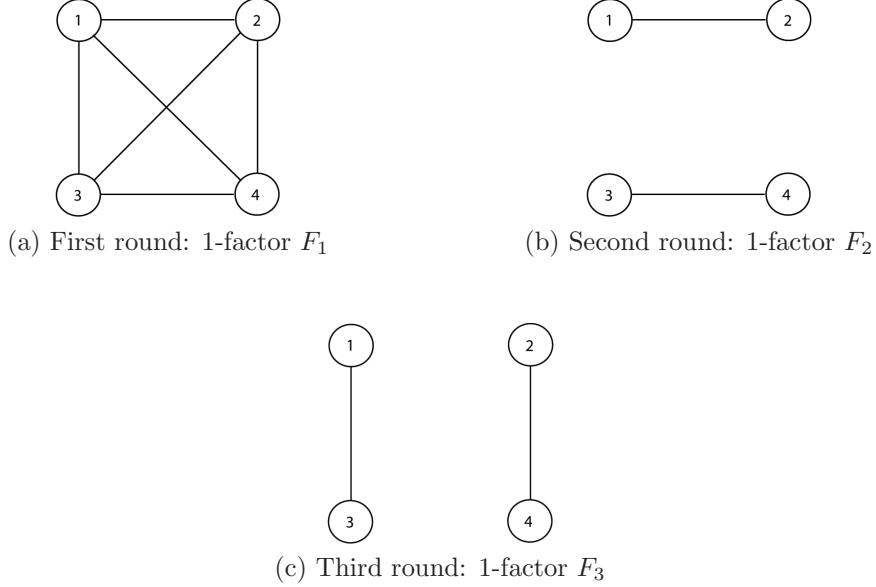


FIGURE 2. Timetable of a single round robin tournament with $n = 4$ teams represented one of its 1-factorizations: (a) 1-factor F_1 associated with the first round, (b) 1-factor F_2 associated with the second round, and (c) 1-factor F_3 associated with the third round.

Some round robin scheduling problems involve the construction of both the timetable and the home-away pattern set. However, either the timetable or the HAP set may be predefined and known beforehand in some situations. In the first case, the timetable is given and the problem consists in finding a feasible HAP set optimizing a certain objective function. In the second case, the home-away pattern set is predetermined and a timetable is requested. The problem of constructing a timetable compatible with a given HAP set and optimizing a certain objective appears as a subproblem in several approaches to solve real-life scheduling problems, see e.g. Nemhauser and Trick (1998).

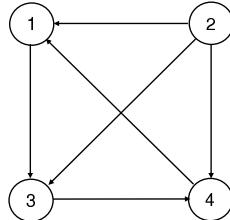


FIGURE 3. Example of a single round robin tournament of four teams represented by an oriented graph in which an arc from node i to j means that team i plays away against team j .

Rounds	Teams			
	1	2	3	4
1	home	away	home	away
2	away	away	home	home
3	home	away	home	away

TABLE 1. Example of a HAP set.

For any given schedule or home-away pattern set, we say that there is a home (resp. away) break in round k whenever a teams plays two consecutive games at home (resp. away) in rounds $k - 1$ and k .

In the example depicted in Figures 1 and 2, with the orientations established in Figure 3, team 1 has a perfectly alternating schedule with no breaks. Team 2 has two away breaks in the second and third rounds, while teams 3 and 4 have one home break each, respectively in the third and second rounds. Break minimization problems deal with the minimization of the number of breaks in the schedule, while distance minimization problems call for the minimization of the total distance traveled by the teams.

3. SOME FUNDAMENTAL PROBLEMS

This section presents an introduction to the main problems in sports scheduling and their variants: breaks minimization, distance minimization and the traveling tournament problem, carry-over effects minimization, and balanced tournament designs.

3.1. Breaks minimization. One of the most important goals in sports scheduling is the minimization of breaks. League organizers seek schedules with a minimum number of breaks or, at least, with a balanced number of breaks (i.e., schedules in which all teams have the same number of breaks). Pioneering work in the area was developed by de Werra (1980; 1981; 1982; 1988).

Sports scheduling problems are often solved by either one of two decomposition approaches:

- (1) “First-schedule, then-break”: first determine the games to be played at each round, next the corresponding home-away pattern.
- (2) “First-break, then-schedule”: first determine a feasible home-away pattern set, next the corresponding games to be played at each round.

Both approaches have been studied in the literature for different problem settings. In both cases, a home-away pattern with a minimum number of breaks is often sought. Break minimization issues in this context have been discussed by Brouwer et al. (2008), Miyashiro et al. (2003), Miyashiro and Matsui (2005), Post and Woeginger (2006), and de Werra et al. (1990). Results on the number of breaks in round robin tournaments also appeared e.g. in (Schreuder, 1980; 1992; de Werra, 1980; 1981; 1982; 1988). Optimization and constraint programming approaches for break minimization have been presented by Regin (2001) and Rasmussen and Trick (2007).

Urrutia and Ribeiro (2006) established a relationship between two aspects of round robin tournament scheduling problems: trips and breaks. In particular, they have shown that, for any schedule S , the total number of travels (or trips) $T(S)$ and the total number of breaks $B(S)$ are such that $T(S) = n \cdot R - B(S)/2$, where

$R = n - 1$ (resp. $R = 2(n - 1)$) is the number of rounds in a compact single (resp. double) round robin tournament with n teams. This connection between breaks maximization and distance minimization was used to derive lower bounds for some instances of the mirrored traveling tournament problem and to prove the optimality of solutions found by a heuristic for the latter.

3.2. Distance minimization and the traveling tournament problem. In the case of distance minimization scheduling problems, there is a distance (or a time or a cost) associated to each pair of teams, corresponding to the traveling distance between their home cities. A schedule minimizing the total distance traveled by all teams is sought. Additional constraints are usually imposed on traveling.

The traveling tournament problem (TTP) introduced in the seminal paper of Easton et al. (2001) is by far the most emblematic problem in this area. It is a challenging combinatorial optimization problem in sports scheduling that abstracts the most important aspects in creating timetables whenever traveling distances are an important issue. Given an even number n of teams, distances d_{ij} between the home cities of teams i and j , for every $i, j = 1, \dots, n$ (with $d_{ii} = 0$ if $i = j$), and two integer numbers L and U , the TTP calls for the schedule of a double round robin tournament minimizing the total distance traveled by the teams and respecting a set of constraints, while assuming that whenever a team plays two consecutive away games, it goes directly from the site of the first opponent to that of the second:

- every team begins the tournament at home and must return to home after its last away game;
- no repeaters are allowed, i.e. no two teams can play against each other in two consecutive rounds;
- every sequence of consecutive home games played by any team is formed by at least L and at most U games; and
- every sequence of consecutive away games played by any team is formed by at least L and at most U games.

The most direct integer programming formulation of the traveling tournament problem makes use of the following decision variables:

$$x_{ijk} = \begin{cases} 1, & \text{if team } i \text{ plays away against team } j \text{ in round } k, \\ 0, & \text{otherwise;} \end{cases}$$

and

$$y_{tijk} = \begin{cases} 1, & \text{if team } t \text{ travels from the facility of team } i \text{ to that of} \\ & \text{team } j \text{ between rounds } k \text{ and } k + 1, \\ 0, & \text{otherwise.} \end{cases}$$

Using these variables, we obtain the formulation (1) to (12) for the TTP:

$$(1) \quad \min \sum_{i=1}^n \sum_{j=1}^n d_{ij} \cdot x_{ij1} + \sum_{t=1}^n \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^{2n-3} d_{ij} \cdot y_{tijk} + \sum_{i=1}^n \sum_{j=1}^n d_{ji} \cdot x_{ij,2n-2}$$

subject to:

$$(2) \quad x_{iik} = 0, \quad i = 1, \dots, n, \quad k = 1, \dots, 2n - 2,$$

$$(3) \quad \sum_{j=1}^n (x_{ijk} + x_{jik}) = 1, \quad i = 1, \dots, n, \quad k = 1, \dots, 2n - 2,$$

$$(4) \quad \sum_{k=1}^{2n-2} x_{ijk} = 1, \quad i, j = 1, \dots, n : i \neq j,$$

$$(5) \quad L \leq \sum_{\ell=0}^U \sum_{j=1}^n x_{ij,k+\ell} \leq U, \quad i = 1, \dots, n, \quad k = 1, \dots, 2n - 2 - U,$$

$$(6) \quad x_{ijk} + x_{jik} + x_{ij,k+1} + x_{ji,k+1} \leq 1, \quad i, j = 1, \dots, n, \quad k = 1, \dots, 2n - 3,$$

$$(7) \quad z_{iik} = \sum_{j=1}^n x_{jik}, \quad i = 1, \dots, n, \quad k = 1, \dots, 2n - 2,$$

$$(8) \quad z_{ijk} = x_{ijk}, \quad i, j = 1, \dots, n : i \neq j, \quad k = 1, \dots, 2n - 2,$$

$$(9) \quad y_{tijk} \geq z_{tik} + z_{tj,k+1} - 1, \quad t, i, j = 1, \dots, n, \quad k = 1, \dots, 2n - 3,$$

$$(10) \quad x_{ijk} \in \{0, 1\}, \quad i, j = 1, \dots, n, \quad k = 1, \dots, 2n - 2,$$

$$(11) \quad z_{ijk} \in \{0, 1\}, \quad i, j = 1, \dots, n, \quad k = 1, \dots, 2n - 2,$$

$$(12) \quad y_{tijk} \in \{0, 1\}, \quad t, i, j = 1, \dots, n, \quad k = 1, \dots, 2n - 2.$$

The objective function (1) accounts for the total traveled distance separated in three terms: the distance traveled by the teams that play away in the first round, the distance traveled after the first and before the last game played by each team, and the distance traveled to return home by the teams that play away in the last round. Constraints (2) state that no team plays against itself, while constraints (3) enforce that each team plays exactly once in each round, either at home or away. Constraints (4) guarantee that each team will play away against each opponent exactly once. Constraints (5) are used to enforce that in any sequence of $U + 1$ consecutive games a team will play at least L and at most U away games. The non-occurrence of repeaters is ensured by constraints (6). Constraints (7) enforce z_{iik} to be equal to 1 (resp. 0) if team i plays at home (resp. away) in round k . For any two different teams i and j , constraints (8) enforce $z_{ijk} = x_{ijk}$, i.e. team i should be at the home city of team j if the former plays away against the latter. Constraints (9) impose team t to travel from the home city of team i to that of team j if it plays in such cities in two consecutive rounds. Constraints (10) to (12) impose the integrality requirements.

Although the above gives a complete formulation for the TTP, the lower bounds provided by its linear programming relaxation are very weak. To improve this formulation, Trick (2003) suggested adding the so called odd-set constraints for each week. An alternative (and much better) approach is to reformulate by redefining the decision variables, as described in (Trick, 2005). We shall return to the issue of problem reformulation in Section 4, where alternative formulations for a variant of the traveling tournament problem will be explored.

The mirrored traveling tournament problem (Ribeiro and Urrutia, 2007b) and the traveling tournament problem with predefined venues (Costa et al., to appear;

Melo et al., 2009) are two variants of the traveling tournament problem. The first has the additional constraint that games played in round t are exactly the same played in round $t + (n - 1)$ for $t = 1, \dots, n - 1$, but with reversed venues. The second is a single round robin variant of the TTP, in which the venue of each game to be played is known beforehand.

Benchmark instances and their best lower and upper bounds for the widely studied case of the TTP with $L = 1$ and $U = 3$ can be found in (Trick, 2010).

The TTP and its variants have been tackled by different exact and approximate solution methods. The first integer programming approach for exactly solving the TTP was proposed by Easton et al. (2003), where the so-called independent lower bound later improved by Urrutia et al. (2007) was originally presented. Rasmussen and Trick (2006) developed an exact two-phase hybrid approach which generates all feasible patterns in a first phase using constraint programming and assigns teams to patterns in the second phase using integer programming. Mirrored and non-mirrored benchmark TTP instances with eight teams have been solved to optimality by Cheung (2008) and Irnich (2010). Uthus et al. (2011) developed an iterative-deepening A* based approach for the traveling tournament problem that was able to find optimal solutions to the largest benchmark instances solved to date, involving ten teams.

Metaheuristics are among the most effective solution strategies for solving combinatorial optimization problems in practice and have been largely applied in the solution of the TTP and its variants. Among the main algorithmic contributions, we cite the hybrid algorithms proposed by Anagnostopoulos et al. (2003; 2006) for the TTP, based on simulated annealing and exploring both feasible and infeasible schedules, and by Ribeiro and Urrutia (2007b) for the mirrored TTP, in which components borrowed from the GRASP and ILS metaheuristics are combined and an ejection-chain mechanism is used to generate perturbations. Numerical results for the mirrored variant have been later improved by Van Hentenryck and Vergados (2006), extending their previous work developed for the general case.

Bhattacharyya (2009) gave the first NP-completeness proof for the variant of the TTP where no constraints exist on the number of consecutive home games or away games of a team. Later, Thielen and Westphal (2011) have shown that the original TTP is strongly NP-complete when the upper bound on the maximal number of consecutive away games is set to 3.

3.3. Carry-over effects minimization. A major issue in the strategy of teams or athletes, in particular in long competitions, consists in balancing their efforts over the competition. If a team plays against a weak opponent, it is likely to be in better shape to play in the next round than if it had played against a hard opponent before. Teams that play against strong opponents will very likely be more tired for their next game. Therefore, it is likely that a team (or an athlete) makes much less effort playing against an opponent that played before against a very strong contestant, than it would make against an opponent that faced an easy contestant.

The above situation is particularly true in the case of sports which require a great amount of physical effort (such as wrestling, rugby, and martial arts). In this sort of sports, it is not uncommon that a team (or an athlete) plays several matches in a row, making a sequence of very tired (resp. well reposed) opponents very attractive (resp. unattractive). Some schedules may contain several of such sequences of easier or harder games assigned to one or more teams. This situation

does not characterize a fair schedule and is highly undesirable in any tournament. To illustrate this effect, suppose a Karate-Do or Judo competition, for which there is no weight division in open-weight categories: a physically weak athlete may fight a strong one. A contestant that has just fought a very strong opponent will possibly be very tired (and often wounded) in his/her next fight. This would deteriorate his/her performance, giving to the next opponent a strong advantage that otherwise he/she would not have.

Although some authors advocate that carry-over effects do not play a major role in collective sports (Goossens and Spieksma, 2009b), Flatberg et al. (2009) have shown a real-life application to a football league in Norway in which carry-over effects determined by one specific team and player strongly affected the final results of the competition. Furthermore, they have also shown that the minimization of such effects lead to a more fair fixture and better schedule of games. Another interesting real-life situation is illustrated by problems in US college football, whereby a team (Alabama) was repeatedly scheduled against teams with byes in the week before. The sequence of games was very unattractive for Alabama, because it was supposed to often meet a restful team that has not played in the previous round (Goodbread, 2010).

For a given compact single round robin schedule with n teams, we say that team i gives a carry-over effect to team j if some other team plays in consecutive rounds against teams i and j . Rounds are considered cyclically, i.e., the first round follows the last round $n - 1$ and may be considered as round n . If team i is a very strong (resp. very weak) team and several teams play consecutively against teams i and j , then team j may be favored (resp. handicapped) when compared with other teams.

Let $c_{ij} \geq 0$ count the number of carry-over effects given by team i to team j , for any $i, j = 1, \dots, n$ with $i \neq j$. The quality of a schedule with respect to carry-over effects is measured by the carry-over effects value $\sum_{i=1}^n \sum_{j=1}^n c_{ij}^2$. In an ideal or balanced schedule with respect to carry-over effects, no teams i, j, p, q should exist such that teams p and q both play against team j immediately after playing against team i . In that case, one should have $c_{ij} = 1$, for any $i, j = 1, \dots, n$ with $i \neq j$.

For every pair of teams $i, j = 1, \dots, n$ (with $i \neq j$) and for every round $k = 1, \dots, n$ (with rounds cyclically represented, such as that round $n - 1$ is followed by the first round), we define the binary variable

$$y_{kij} = \begin{cases} 1, & \text{if team } i \text{ plays against team } j \text{ in round } k, \\ 0, & \text{otherwise.} \end{cases}$$

The number of carry-over effects given by team i to j is $\sum_{\ell=1}^n \sum_{k=1}^{n-1} y_{k\ell i} \cdot y_{(k+1)\ell j}$, for any $i \neq j$; zero otherwise. Therefore, the carry-over effects value minimization problem (13) to (19) can be formulated by integer programming as follows:

$$(13) \quad \min \sum_{i=1}^n \sum_{j=1}^n \left(\sum_{\ell=1}^n \sum_{k=1}^{n-1} y_{k\ell i} \cdot y_{(k+1)\ell j} \right)^2$$

subject to:

$$(14) \quad y_{kij} = y_{kji}, \quad i, j, k = 1, \dots, n,$$

$$(15) \quad \sum_{i=1}^n y_{kij} = 1, \quad j, k = 1, \dots, n,$$

$$(16) \quad \sum_{k=1}^{n-1} y_{kij} = 1, \quad i, j = 1, \dots, n : i \neq j,$$

$$(17) \quad y_{kii} = 0, \quad i, k = 1, \dots, n,$$

$$(18) \quad y_{nij} = y_{1ij}, \quad i, j = 1, \dots, n,$$

$$(19) \quad y_{kij} \in \{0, 1\}, \quad i, j, k = 1, \dots, n.$$

The objective function (13) minimizes the total carry-over effects value. Constraints (14) ensure that variables $y_{kij} = y_{kji}$ are the same or, alternatively, that the game between teams i and j is the same as that between teams j and i (there is a unique game between teams i and j in a single round robin tournament). Constraints (15) enforce that every team plays exactly once in each round of a compact schedule. Constraints (16) and (17) guarantee that each team plays exactly once against every other team. Constraints (18) enforce that the n -th round is equivalent to the first. Constraints (19) impose the binary requirements on the variables. This formulation has $O(n^3)$ variables and defines a quadratic minimization problem. The linearization of the objective function (13) leads to a reformulation with $O(n^4)$ variables.

Russell (1980) proposed a construction algorithm that generates schedules matching the lower bound to the carry-over effects value when the number of teams is a power of two. The method proposed by Anderson (1999) obtained solutions that are still the best known to date (except for $n = 12$). It makes use of algebraic structures called starters (Dinitz, 1996) to generate schedules. However, the approach presumes that a suitable starter is known beforehand, which may imply in huge computation times.

Trick (2000) developed a constraint programming method that made it possible to prove the optimality of Russell's method for $n = 6$. Henz et al. (2004) improved the solution obtained by the previous approach for $n = 12$, also using constraint programming. Miyashiro and Matsui (2006) developed a time-consuming heuristic based on random permutations of the rounds of fixtures created by the polygon method (Kirkman, 1847). They reported more than two days of computation times for $n \geq 18$.

Guedes and Ribeiro (2009) developed an ILS-based heuristic for solving a weighted variant of the carry-over effects minimization problem. This heuristic obtained the best known solution at the time of writing for the unweighted instance with $n = 12$.

4. REFORMULATIONS

Straightforward problem formulations often lead to weak linear relaxation bounds to integer programs. Reformulation techniques based on the redefinition of the decision variables used to formulate the problem provide an effective strategy to obtain tighter relaxations with improved bounds. Trick (2005) explored problem reformulation in the context of the traveling tournament problem. In this paper, we consider the traveling tournament problem with predefined venues to illustrate this approach.

Melo et al. (2009) have introduced the traveling tournament problem with predefined venues (TTPPV), as already discussed in Section 3.2. This problem is a single round robin variant of the TTP, in which the venues where the games take place are known beforehand. Variants of this problem find interesting applications in real-life leagues whose DRR tournaments are divided into two SRR phases. Games in the second phase are exactly the same of the first phase, except for the inversion of their venues. Therefore, the venues of the games in the second phase are known beforehand and constrained by those of the games in the first phase. This is the case e.g. of the Chilean soccer professional league (Durán et al., 2007b) and of the German table tennis federation of Lower Saxony (Knust, 2007). As before, we assume the tournament is played by an even number n of teams, indexed by $1, \dots, n$. Each team has its own venue at its home city. All teams are initially at their home cities, to where they must return after their last away game. The distance $d_{ij} \geq 0$ from the home city of team i to that of team j is known, for every $i, j = 1, \dots, n$, with $i \neq j$. A road trip is a sequence of consecutive away games played by a team at the venues of its opponents, along which this team travels from the venue of one opponent to that of the next, without returning home. As for the most studied case of the TTP, we assume that every sequence of consecutive home (or away) games played by any team is formed by at least $L = 1$ and at most $U = 3$ games.

Let G be a set of games, represented by ordered pairs of teams. The game between teams i and j is represented either by the ordered pair (i, j) or by the ordered pair (j, i) . In the first case, the game between i and j takes place at the venue of team i ; otherwise, at that of team j . For every two teams i and j , either $(i, j) \in G$ or $(j, i) \in G$. The problem consists therefore in finding a compact single round robin schedule compatible with G , such that the total distance traveled by the teams is minimized and no team plays more than three consecutive home games or three consecutive away games. Melo et al. (2009) proposed three integer programming formulations for this problem, using $O(n^3)$, $O(n^4)$, and $O(n^5)$ variables. We describe and compare in the following the two formulations in $O(n^3)$ and $O(n^5)$ variables.

4.1. Formulation with $O(n^3)$ variables. We define two types of decision variables:

$$z_{tjk} = \begin{cases} 1, & \text{if team } t \text{ plays at home against team } j \text{ in round } k, \\ 0, & \text{otherwise;} \end{cases}$$

and

$$y_{tij} = \begin{cases} 1, & \text{if team } t \text{ travels from the facility of team } i \text{ to that of team } j, \\ 0, & \text{otherwise.} \end{cases}$$

The y variables represent the journeys performed by a team between pairs of cities. Since each game occurs exactly once, a journey between the home cities of two different teams is performed at most once by each team. Variables z and y are used in the formulation (20) to (32) of TTPPV, which has $O(n^3)$ variables:

$$(20) \quad \min F_1(z, y) = \sum_{t=1}^n \sum_{i=1}^n \sum_{j=1}^n d_{ij} \cdot y_{tij}$$

subject to:

$$(21) \quad \sum_{q=1}^{n-1} z_{tjq} = 1, \quad \forall (t, j) \in G,$$

$$(22) \quad \sum_{q=1}^{n-1} z_{tjq} = 0, \quad \forall (j, t) \in G,$$

$$(23) \quad \sum_{\substack{j=1 \\ j \neq t}}^n (z_{tjk} + z_{jtk}) = 1, \quad t = 1, \dots, n, \quad k = 1, \dots, n-1,$$

$$(24) \quad y_{tij} \geq z_{it,k-1} + z_{jtk} - 1, \quad t, i, j = 1, \dots, n : t \neq i \neq j, \quad k = 2, \dots, n-1,$$

$$(25) \quad y_{tit} \geq z_{it,k-1} + \sum_{\substack{j=1 \\ j \neq t}}^n z_{tjk} - 1, \quad t, i = 1, \dots, n : t \neq i, \quad k = 2, \dots, n-1,$$

$$(26) \quad y_{tti} \geq \sum_{\substack{j=1 \\ j \neq t}}^n z_{tj,k-1} + z_{itk} - 1, \quad t, i = 1, \dots, n : t \neq i, \quad k = 2, \dots, n-1,$$

$$(27) \quad y_{tti} \geq z_{it1}, \quad t, i = 1, \dots, n : t \neq i,$$

$$(28) \quad y_{tit} \geq z_{it,n-1}, \quad t, i = 1, \dots, n : t \neq i,$$

$$(29) \quad \sum_{q=k}^{k+3} \sum_{\substack{j=1 \\ j \neq t}}^n z_{jqt} \leq 3, \quad t = 1, \dots, n, \quad k = 1, \dots, n-4,$$

$$(30) \quad \sum_{q=k}^{k+3} \sum_{\substack{j=1 \\ j \neq t}}^n z_{jqt} \geq 1, \quad t = 1, \dots, n, \quad k = 1, \dots, n-4,$$

$$(31) \quad z_{tjk} \in \{0, 1\}, \quad t, j = 1, \dots, n, \quad k = 1, \dots, n-1,$$

$$(32) \quad y_{tij} \in \{0, 1\}, \quad t, i, j = 1, \dots, n.$$

The objective function (20) defines the minimization of the total distance traveled by the teams. Constraints (21) and (22) ensure that each game in G occurs exactly once. Constraints (23) guarantee that each team plays one game in each round. Constraints (24) enforce team t to perform a trip from the home city of team i to that of team j if it plays two consecutive away games against teams i and j , in this order. Constraints (25) enforce team t to perform a trip from the home city of team i to its own home city if it has an away game against the latter followed by a home game in the next round. Constraints (26) enforce team t to travel from its own home city to that of team i to play away against the latter after a home game in the previous round. Constraints (27) enforce team t to travel to the home city of team i if it plays away against the latter in the first round. Constraints (28) enforce team t to return from the home city of team i if it plays away against the latter in the last round. Together, constraints (29) and (30) imply that team t cannot play less

than one or more than three home games (or away games) in any sequence of four consecutive games. Constraints (31) and (32) define the integrality requirements.

This formulation has $O(n^4)$ constraints: $O(n^2)$ of types (21), (23), (27), (28), (29) and (30), $O(n^3)$ of types (25) and (26), and $O(n^4)$ of type (24).

4.2. Formulation with $O(n^5)$ variables. This formulation considers complete road trips. Its variables represent road trips of different sizes, giving a more direct representation of the problem. Three new types of decision variables are defined and used in this formulation:

$$w_{tik}^1 = \begin{cases} 1, & \text{if team } t \text{ starts, in round } k, \text{ a road trip visiting team } i \text{ and} \\ & \text{returning home in round } k+1 \text{ (with } t \neq i\text{),} \\ 0, & \text{otherwise;} \end{cases}$$

$$w_{tijk}^2 = \begin{cases} 1, & \text{if team } t \text{ starts, in round } k, \text{ a road trip visiting first team } i, \text{ then} \\ & \text{team } j, \text{ and returning home in round } k+2 \text{ (with } t \neq i \neq j\text{),} \\ 0, & \text{otherwise;} \end{cases}$$

and

$$w_{tij\ell k}^3 = \begin{cases} 1, & \text{if team } t \text{ starts, in round } k, \text{ a road trip visiting first team } i, \text{ then} \\ & \text{team } j, \text{ next team } \ell, \text{ and returning home in round } k+3 \\ & (\text{with } t \neq i \neq j \neq \ell), \\ 0, & \text{otherwise.} \end{cases}$$

Two dummy rounds (indexed by -1 and 0) are created to simplify the formulation. The variables corresponding to every road trip starting in any of these dummy rounds are set to 0. The auxiliary costs c_{ij} , c_{ijm} , and c_{ijml} represent the costs of road trips of length one, two, and three performed by team i , respectively:

$$\begin{aligned} c_{ij} &= d_{ij} + d_{ji}, \\ c_{ijm} &= d_{ij} + d_{jm} + d_{mi}, \text{ and} \\ c_{ijml} &= d_{ij} + d_{jm} + d_{ml} + d_{\ell i}. \end{aligned}$$

The new variables are used to reformulate TPPV as the integer program (33) to (41) below, which has $O(n^5)$ variables:

$$(33) \quad \min F_3(w^1, w^2, w^3) = \sum_{k=1}^{n-1} \sum_{i=1}^n \sum_{\substack{j=1 \\ (j,i) \in G}}^n [c_{ij} \cdot w_{tijk}^1 + \sum_{\substack{m=1 \\ (m,i) \in G \\ m \neq j}}^n (c_{ijm} \cdot w_{tijk}^2 + \sum_{\substack{\ell=1 \\ (\ell,i) \in G \\ \ell \neq j \neq m}}^n c_{ijml} \cdot w_{tijk\ell}^3)]$$

subject to:

$$(34) \quad \sum_{i=1}^n \sum_{\substack{j=1 \\ (j,i) \in G}}^n [\sum_{k \in \{-1,0\}} w_{tijk}^1 + \sum_{\substack{m=1 \\ (m,i) \in G \\ m \neq j}}^n (\sum_{k \in \{-1,0,n-1\}} w_{tijk}^2 + \sum_{\substack{\ell=1 \\ (\ell,i) \in G \\ \ell \neq j \neq m}}^n w_{tijk\ell}^3)] = 0,$$

$$(35) \quad \sum_{k=1}^{n-1} \{w_{ijk}^1 + \sum_{\substack{m=1 \\ (m,i) \in G \\ m \neq j}}^n [(w_{ijmk}^2 + w_{imjk}^2) + \sum_{\substack{\ell=1 \\ (\ell,i) \in G \\ \ell \neq j \neq m}}^n (w_{ijm\ell k}^3 + w_{imj\ell k}^3 + w_{im\ell jk}^3)]\} = 1, \quad \forall (j,i) \in G,$$

$$(36) \quad \sum_{\substack{j=1 \\ (j,i) \in G}}^n \{w_{ijk}^1 + \sum_{\substack{m=1 \\ (m,i) \in G \\ m \neq j}}^n [(w_{ijmk}^2 + w_{ijm,k-1}^2) + \sum_{\substack{\ell=1 \\ (\ell,i) \in G \\ \ell \neq j \neq m}}^n (w_{ijm\ell k}^3 + w_{ijm\ell,k-1}^3 + w_{ijm\ell,k-2}^3)]\} + \sum_{\substack{j=1 \\ (i,j) \in G}}^n \{w_{jik}^1 + \sum_{\substack{m=1 \\ (m,j) \in G \\ m \neq i}}^n [(w_{jimk}^2 + w_{jmi,k-1}^2) + \sum_{\substack{\ell=1 \\ (\ell,j) \in G \\ \ell \neq i \neq m}}^n (w_{jim\ell k}^3 + w_{jmi\ell,k-1}^3 + w_{jmi\ell,k-2}^3)]\} = 1, \quad i = 1, \dots, n, \quad k = 1, \dots, n-1,$$

$$(37) \quad \sum_{\substack{j=1 \\ (j,i) \in G}}^n \{w_{ijk}^1 + w_{ij,k+1}^1 + \sum_{\substack{m=1 \\ (m,i) \in G \\ m \neq j}}^n [w_{ijm,k-1}^2 + w_{ijmk}^2 + w_{ijm,k+1}^2 + \sum_{\substack{\ell=1 \\ (\ell,i) \in G \\ \ell \neq j \neq m}}^n (w_{ijm\ell,k-2}^3 + w_{ijm\ell,k-1}^3 + w_{ijm\ell k}^3 + w_{ijm\ell,k+1}^3)]\} \leq 1, \quad i = 1, \dots, n, \quad k = 1, \dots, n-2,$$

$$(38) \quad \sum_{q=k}^{k+3} \{ \sum_{\substack{j=1 \\ (j,i) \in G}}^n [w_{ijq}^1 + \sum_{\substack{m=1 \\ (m,i) \in G \\ m \neq j}}^n ((w_{ijmq}^2 + w_{ijm,q-1}^2) + \sum_{\substack{\ell=1 \\ (\ell,i) \in G \\ \ell \neq j \neq m}}^n (w_{ijm\ell q}^3 + w_{ijm\ell,q-1}^3 + w_{ijm\ell,q-2}^3))] \} \geq 1, \quad i = 1, \dots, n, \quad k = 1, \dots, n-4,$$

$$(39) \quad w_{ijk}^1 \in \{0, 1\}, \quad i, j = 1, \dots, n : i \neq j, \quad k = -1, \dots, n-1,$$

$$(40) \quad w_{ijmk}^2 \in \{0, 1\}, \quad i, j, m = 1, \dots, n : i \neq j \neq m, \quad k = -1, \dots, n-1,$$

$$(41) \quad w_{ijm\ell k}^3 \in \{0, 1\}, \quad i, j, m, l = 1, \dots, n : i \neq j \neq m \neq \ell, \quad k = -1, \dots, n-1.$$

The objective function (33) minimizes the total traveled distance. Constraints (34) set to zero the variables associated to road trips starting at the dummy rounds -1 and 0, to road trips of size two and three starting at the last round and those of size three starting at round $n-2$. Constraints (35) ensure that each game occurs exactly once. They represent the fact that each game $(j,i) \in G$ should be played in a road trip of team i formed by one, two, or three away games. Constraints (36) enforce that team i is either playing an away game or another team is visiting it in

each round. This is achieved by setting to one the sum of all variables associated with road trips of team i which include round k and road trips of other teams which visit i in round k . Constraints (37) forbid team i to be engaged in simultaneous or consecutive (i.e., without returning to its home city) road trips in round k . Constraints (38) state that team i must be outside its home city to play away at least once along every four consecutive rounds. Constraints (39) to (41) guarantee the integrality requirements.

Although the number of variables increases with respect to the previous formulation, we notice that the number of constraints is quite smaller: this formulation has $O(n^2)$ constraints of types (35) to (38).

4.3. Bounds. Theorems 1 and 2 in (Melo et al., 2009) prove that the bound $LB5$ provided by the second formulation, with $O(n^5)$ variables, is at least as good as the bound $LB3$ given by the first, with $O(n^3)$ variables.

To further evaluate and compare the numerical results obtained by the two integer programming formulations, Melo et al. (2009) derived ten test instances from each national league (NL) and each circular (CIRC) benchmark TTP instance available in (Trick, 2010) and applied CPLEX 10.0 to solve them. The computational experiments were performed on a Dell Optiplex machine, with a Pentium D 3.0 GHz processor and 2 GB of RAM memory.

TABLE 2. Linear relaxation lower bounds for random CIRC instances

n	Feasible instances	avg. $LB3$	avg. $LB5$	average value of $(LB5 - LB3)/LB3 (\%)$	maximum value of $(LB5 - LB3)/LB3 (\%)$
4	10	4.94	12.60	154.50	180.00
6	7	8.50	35.90	329.52	410.81
8	8	14.53	73.78	431.45	702.81
10	7	12.81	132.76	971.62	1197.72
12	10	15.03	214.23	1402.89	1667.04
14	8	16.35	326.62	1907.82	2179.34
16	10	18.90	471.70	2488.63	2860.97
18	8	19.24	649.38	3294.49	3528.10
20	8	20.83	869.29	4102.11	4252.44

The linear relaxations of the two formulations have been solved for all instances. Since the results obtained for the different types of test problems are very similar, we report in Table 2 the results obtained exclusively for the CIRC instances with random assignments of venues. For each value of n , the second column of this table displays the number of feasible instances (out of a total of ten). The third and fourth columns give, respectively, the average lower bounds $LB3$ and $LB5$ for the feasible instances. The two last columns give the average and maximum gaps between the bounds $LB3$ and $LB5$, with the relative gap computed as $100 \cdot (LB5 - LB3)/LB3$ informing by how much $LB5$ improves upon $LB3$. The average value of the lower bound $LB5$ is far better than that of $LB3$. Considering e.g. the CIRC instances with $n = 20$, the average bound $LB5$ is approximately 41 times greater than the average bound $LB3$ for the random instances.

5. APPLICATIONS

In this section, we survey applications of optimization methods to scheduling problems in different sport disciplines such as football, baseball, basketball, cricket, and hockey. Basketball and football are the sports with more applications.

5.1. Football. In spite of the large number of papers reporting on applications of integer programming and metaheuristics to the schedule of football tournaments, Nurmi et al. (2010) noticed that only a few refer to real-life applications funded by research and development contracts. Among them, Bartsch et al. (2006) applied heuristics and branch-and-bound to schedule the professional football leagues of Austria and Germany. Della Croce and Oliveri (2006) adapted the integer programming approach of Nemhauser and Trick (1998) to schedule the Italian football league. The proposed procedure is divided into three phases. The first phase generates a pattern set respecting the cable televisions requirements and several other constraints. The second produces a feasible round robin schedule compatible with the above pattern set. The third phase generates the actual calendar, assigning teams to patterns.

Durán et al. (2007a; 2009) also used integer programming to schedule the Chilean football league. Improvements to their solution approach have been proposed by Durán et al. (2007b). Goossens and Spieksma (2009a) reported on the application of integer programming to schedule the Belgian football league for the seasons 2006-2007 and 2007-2008. Rasmussen (2008) presented a solution approach using a logic-based Benders decomposition and column generation to solve a triple round robin tournament for the Danish football league.

Ribeiro and Urrutia (2007a; 2009; 2010) developed an integer programming decomposition strategy to solve the bicriteria optimization problem arising from the schedule of the Brazilian national football tournament, in which one of the objectives consists in maximizing the number of games that may be broadcast by open TV channels (to increase the revenues from broadcast rights) and the other consists in finding a balanced schedule with a minimum number of home breaks and away breaks (for sake of fairness). Their approach has been used in practice to schedule the 2009, 2010, and 2011 editions of the tournament (Ribeiro and Urrutia, 2011). Fiallos et al. (2010) developed an integer programming model solvable by CPLEX that was used for the first time in 2010 to schedule the double round robin professional football tournament of Honduras, played by ten teams.

5.2. Basketball. Campbell and Chen (1976) have been the first to consider the problem of scheduling a basketball conference of ten teams, corresponding to a relaxed double round robin tournament. The teams were allowed to play at most two consecutive away games without returning home. In the first phase, optimal trips for each team were derived. This was shown to be equivalent to pairing the teams two by two, such that the distances between the paired teams were minimized. In the second phase, the optimal pairing was used to build a number of feasible sequences using a constructive approach attempting to minimize the total traveled distance. Ball and Webster (1977) tackled a similar problem. Travel distances were minimized using an integer programming formulation, which was solved by a heuristic very similar to the method developed in the previous reference.

For the National Basketball Association (NBA) in the United States, Bean and Birge (1980) constructed schedules for 22 teams where each team played 82 games

and resting times and building availabilities had to be taken into account. Methods based on heuristics for the traveling salesman problem were proposed in order to reduce the airline traveling costs.

Nemhauser and Trick (1998) combined integer programming and enumerative techniques for determining the schedule of games of the nine universities in the Atlantic Coast Conference (ACC). The problem involved many conflicting requirements and preferences. The solution procedure found a schedule that was accepted by the ACC to be played in 1997-1998. Later, Henz (2001) provided a much faster constraint programming approach for the same problem, finding in less than one minute the 179 solutions for which Nemhauser and Trick (1998) reported an overall running time of about 24 hours.

We conclude this section by noticing that Fronček (2001) proposed to construct schedules for the Czech national basketball league using graph models, while Wright (2006) described a real-life multi-objective problem in scheduling basketball tournaments in New Zealand, solved by a variant of simulated annealing and used to produce the 2004 schedule.

It is also interesting to notice that most football tournaments correspond to compact round robin schedules organized in rounds: every team plays exactly once in each round and most rounds are held in weekends. Contrarily, most basketball tournaments have more relaxed schedules, in which the games can take place at any day of the week and the idea of a round is not very well established.

5.3. Cricket. Armstrong and Willis (1993) reported on the first attempt to use optimization methods in the schedule of cricket competitions, addressing the scheduling of the 1992 Cricket World Cup tournament, co-hosted by Australia and New Zealand. Each of the nine teams had to play each other once over two days in the initial stage of the competition. A variety of constraints had to be respected, which included satisfying local populations and worldwide TV audiences, together with other practical and logistical considerations. The solution methodologies were developed using a spreadsheet package. One of the proposed solution methods allowed interaction with the users and was more useful, but took about four hours to run. None of the schedules produced were completely satisfactory. One year later, Willis and Terrill (1994) considered the scheduling of domestic cricket in Australia, including both first class and one day matches. Simulated annealing was used and, after some manual amendments, the schedule was used for the 1992-1993 season. Simulated annealing was also used by Wright (2005) to produce the 2003-2004 schedule for New Zealand cricket and by Wright (1992) in a case study to produce a four-year schedule (1992-1995) for English county cricket.

5.4. Baseball. Computer-aided heuristics have been used to schedule the Major League Baseball clubs by Cain (1977). For the National American Baseball League, schedules for 12 teams divided into two divisions have to be found, where each team played 162 games (18 times against each team of its own division and 12 times against each team of the other division). The objective consisted in determining a schedule regarding fairness aspects, maximizing attendance, and minimizing travel costs. Results have been reported for the seasons 1969 and 1975.

Russell and Leung (1994) devised cost effective schedules for a baseball league. Two heuristics have been presented to enable a low cost schedule to be found. The Texas Baseball League was used as a benchmark example instance.

The research on the traveling tournament problem that lead to the seminal paper of Easton et al. (2001) and the developments that followed have been motivated by the problem of scheduling the Major League Baseball. Its formulation captures the fundamental difficulties involved in minimizing the travel distance for a sports league.

Hoshino and Kawarabayashi (2011a) considered a relaxed (non-compact) variant of the TTP with additional balancing constraints to schedule Japan's biggest and most well-known professional sports league. According to the authors, they have determined a league schedule that meets all conditions and constraints of the 12-team pro baseball league in Japan, while considerably reducing the total travel distance. In addition, Hoshino and Kawarabayashi (2011b;c) have also investigated the bipartite traveling tournament problem, which is an inter-league extension of the TTP. The decision version of the problem is proved to be NP-complete. They proposed heuristics that have been applied to problem instances with data of the Nippon Professional Baseball league in Japan.

5.5. Hockey. Ferland and Fleurent (1991) described a support system to assist the manual creation of schedules for the National Hockey League (NHL), a relaxed tournament with 21 teams. The tournament was divided into two conferences and each conference into two divisions. The schedules were subject to a number of constraints involving aspects such as the places where the games take place, how often teams can play, the minimum time between two games with the same opponents, and the traveling distances. In response to the National Hockey League expanding from 21 teams, Fleurent and Ferland (1993) devised an integer linear programming formulation to investigate how the increase in the number of teams would add to the complexity of generating schedules. The paper considered various scenarios and the solution accepted by the NHL managers for a 24-team problem was shown, being used as the basis for a schedule to which other matches have been manually added. Later, Costa (1995) investigated the hybridization of genetic algorithms with tabu search to solve combinatorial optimization problems, using the National Hockey League as an example to illustrate the effectiveness of the approach.

6. APPLICATION: SCHEDULING THE BRAZILIAN FOOTBALL TOURNAMENT

We describe with more detail in this section a recent application of integer programming in the scheduling of the yearly first division football tournament in Brazil.

The yearly football tournament organized by the Brazilian Football Confederation (CBF) is the most important sport event in the country. Its major sponsor is TV Globo, the largest media group and television network in Brazil. Fair and balanced schedules for all teams are major issues for attractiveness and confidence in the outcome of the tournament. Furthermore, TV sponsors condition their support to schedules that make it possible to broadcast the most important games by open channels. Large cities hosting two or more teams and a large number of fans impose additional security constraints to avoid clashes of fans before or after the games.

The tournament lasts for seven months, from May to December, and is structured as a compact mirrored double round robin tournament played by $n = 20$ teams. There are $2 \cdot (n - 1) = 38$ rounds. Games of *weekend rounds* are played in Saturdays and Sundays, while those of *midweek rounds* are played in Wednesdays

and Thursdays. The dates available for game playing change from year to year and have to be coordinated with other competitions such as regional tournaments, Brazil's Cup, South America's Cup, and Santander Libertadores Cup. In accordance with the definition of a compact mirrored double round robin tournament, if the game between teams A and B is played in the first phase at the venue of A in some round $k = 1, \dots, n - 1$, then the second game between A and B will be played in round k of the second phase (or in the overall round $k + n - 1$), but now at the venue of B. More attractive games and those involving the strongest teams should as much as possible be played on weekends. Teams are organized by pairs with complementary home-away patterns of game playing. Usually, teams in the same pair are based on the same home city. Team pairings are defined by the Brazilian Football Confederation before the construction of the schedule and may differ from year to year, since the participating teams necessarily change due to promotions to and relegations from lower divisions at the end of each season.

The schedule should satisfy a number of hard and soft constraints, ranging from fairness to security issues, and from technical to broadcasting criteria. Most of them reflect strategies for maximizing revenues and tournament attractiveness, while others attempt to avoid unfair situations that could benefit one team or another with a more convenient sequence of games. These requirements have been discussed and established over the years by teams, federations, city administrators, security agencies, and sponsors.

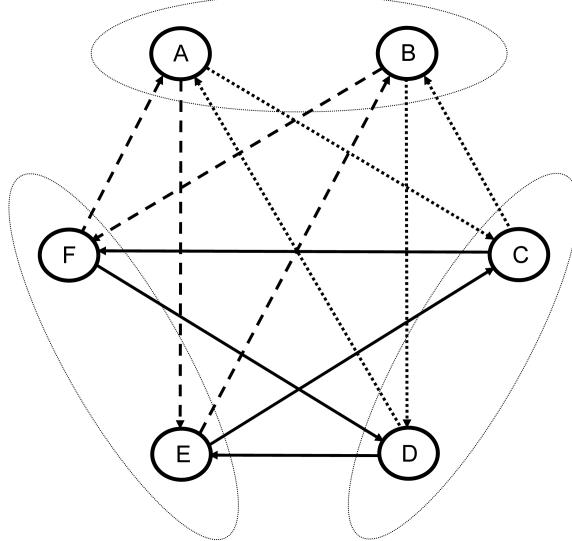
In the quest for a fair and balanced schedule, CBF imposes additional constraints to enforce a tight equilibrium to any two teams belonging to the same pair. Let A and B be two teams belonging to the same pair and C and D two other teams belonging to any other pair. If A plays with C at home (respectively away) in the first phase, then it plays away (respectively at home) with D in this phase. Consequently, B will play away (respectively at home) with C and at home (respectively away) with D in this phase. The same constraints are automatically implied for the second phase, since all its games have interchanged venues with respect to those in the first phase. Such constraints lead to a strong equilibrium between teams in the same cities and regions and are considered by CBF officials as among the most important to be enforced. Figure 4 illustrates a possible perfect matching of six paired teams, which applies to every subset of four teams organized in two pairs.

The maximization of gate attendances and TV audiences is the major issue at stake. Most revenues earned by the teams come from broadcasting and merchandising rights paid by the sponsors, which request good schedules drawing large audiences. Fair and balanced schedules for all teams are also a major issue for the attractiveness of the tournament and for the confidence in its outcome, playing a major role in the success of the competition. To maximize gate attendances and TV audiences, we seek a schedule with a maximum number of attractive games played in weekends.

A detailed description of the problem, together with its integer programming formulation and the decomposition solution approach, can be found in Ribeiro and Urrutia (2010; 2011).

The complete optimization model and the software system coded in C++ have been developed, tuned and validated during 2007 and 2008. Staff of CBF and TV Globo participated actively in the formulation of the problem and in the validation of the results it obtained.

FIGURE 4. In this example of a perfect matching of six teams paired into three pairs, an arrow in the arc from team i to team j means that this game is played at the venue of team j .



In 2009, the system was first used as the official scheduler for Brazilian football tournament. New criteria have been proposed and introduced in the system along the decision process based on successive refinements of the solution, as the decision makers evaluated and filtered the different solutions. The organizers checked each proposed schedule and imposed additional constraints (or removed existing constraints) to handle specific situations that might be desired to fine-tune the schedule. That tournament was the most attractive in recent times, with four teams still in contention for the title when the last round started. All games in the last round started simultaneously. The title changed hands several times, as the scores of the ten games underway changed. The goal that decided the tournament for Flamengo was scored only 20 minutes before the end of the tournament and the champion was not known until the last game ended. This scenario was partly the result of a fair and balanced schedule of games, in which no team had specific advantages or disadvantages.

The optimization system was used for the second time in 2010. Once again, the decision makers were happy with the alternative schedules the system computed and with their choices. This was a particularly difficult tournament to schedule. It had to be interrupted in June and July during the 2010 World Cup; thus, few dates were available for game playing. As a result, there were more midweek rounds and fewer weekend rounds, making it impossible to schedule all classic games (or derbies) in weekend rounds as the organizers originally desired. The system sought a schedule with a maximum number of classic games played at double weekend rounds (i.e., with both matches between the same pair of teams associated with a classic game being ideally played in weekend rounds). As in 2009, the title was decided in the last round, with three teams still in contention for the title when

their matches started. The goal that decided the tournament for Fluminense was scored 25 minutes before the end of the tournament.

The construction of the schedule for the 2011 edition of the tournament placed a new challenge to the system. As an attempt to make the tournament even more thrilling, CBF decided to schedule all classic and most attractive matches between teams with the same home city (local derbies) in the three last rounds of each of its phases. The integer programming model was first used to show that such constraints would make the problem necessarily infeasible. Therefore, the objective function was changed to maximize the number of local derbies that could be played in the last three rounds of each phase, with the secondary objective of maximizing the number of attractive games that could be played in weekends. The solution produced by the system could hardly be found by non-automatic methods and was quickly adopted by CBF.

7. CONCLUDING REMARKS

We have reviewed a number of fundamental problems and formulation issues in sports scheduling, followed by a survey of applications of optimization methods to scheduling problems in professional leagues of different sport disciplines such as football, baseball, basketball, cricket, and hockey.

Although this paper focused mainly on problems and applications related to professional leagues, a number of applications are also reported for amateur leagues. Amateur leagues usually do not have access to the same investments and structure, but tournaments and competitors abound, requiring coordination and logistical efforts.

In the United States, for example, regional amateur leagues of several sports, such as baseball, basketball, and football, have hundreds of games occurring every weekend in different divisions. In a single league in California there might be up to 500 football games in a weekend, to be refereed by hundreds of certified referees. In the MOSA (Monmouth & Ocean Counties Soccer Association) league, New Jersey, boys and girls of ages 8 to 18 make up six divisions per age and gender group with six teams per division, totalizing 396 games every Sunday.

Problems in amateur leagues often have a different nature, due to the diverse interests involved. Knust (2010) and Schönberger et al. (2004) described variants of genetic algorithms for determining schedules of non-professional table-tennis leagues. Duarte et al. (2007a;b) and Duarte and Ribeiro (2008) tackled single and multiobjective versions of a referee assignment problem that are typical of large amateur leagues such as those above mentioned. Referee assignment problems in professional leagues have been addressed by Evans (1988); Evans et al. (1984); Farmer et al. (2007); Ordonez and Knowles (1998); Wright (1991), and Yavuz et al. (2008). In another context, problems of assigning judges to academic competitions have been considered by Lamghari and Ferland (2007; 2010a;b).

The hardness of sports scheduling optimization problems has led to the use of different techniques in their solution. The best results are often obtained by methods derived from the hybridization of integer programming, constraint programming, and metaheuristics.

Devising optimal tournament schedules is crucial to players, teams, fans, cities, security force, TV channels, and other sponsors. Fair and balanced schedules for all teams, satisfying a large number of hard and soft constraints, are a major issue

for the attractiveness and the confidence in the outcome of professional league tournaments.

The success of real-life applications has shown that Operations Research has certainly proved that it has its place in sports management. Besides the quality of the solutions found, the main advantages of optimization-based computational systems for scheduling sports leagues are their ease of use and the construction of several alternative schedules, making it possible for the decision maker planning the competition to compare and select the most attractive schedule from among different alternatives, which can contemplate other secondary goals and constraints.

REFERENCES

- A. Anagnostopoulos, L. Michel, P. Van Hentenryck, and Y. Vergados. A simulated annealing approach to the traveling tournament problem. In *Proceedings of CPAIOR'03*, 2003.
- A. Anagnostopoulos, L. Michel, P. Van Hentenryck, and Y. Vergados. A simulated annealing approach to the traveling tournament problem. *Journal of Scheduling*, 9:177–193, 2006.
- I. Anderson. Balancing carry-over effects in tournaments. In F. Holroyd, K. Quinn, C. Rowley, and B. Webb, editors, *Combinatorial Designs and Their Applications*, CRC Research Notes in Mathematics, pages 1–16. Chapman & Hall, 1999.
- J. Armstrong and R. J. Willis. Scheduling the cricket World Cup – A case-study. *Journal of the Operational Research Society*, 44:1067–1072, 1993.
- B. C. Ball and D. B. Webster. Optimal scheduling for even-numbered team athletic conferences. *AIEE Transactions*, 9:161–169, 1977.
- T. Bartsch, A. Drexl, and S. Kröger. Scheduling the professional soccer leagues of Austria and Germany. *Computers & Operations Research*, 33:1907–1937, 2006.
- J. C. Bean and J. R. Birge. Reducing travelling costs and player fatigue in the National Basketball Association. *Interfaces*, 10:98–102, 1980.
- R. Bhattacharyya. A note on complexity of traveling tournament problem, 2009. Online reference at http://www.optimization-online.org/DB_HTML/2009/12/2480.html, last visited on April 3, 2011.
- A. E. Brouwer, G. Post, and G. J. Woeginger. Tight bounds for break minimization. *Journal of Combinatorial Theory A*, 115:1065–1068, 2008.
- W. O. Cain. The computer-aided heuristic approach used to schedule the Major League Baseball clubs. In S. P. Ladany and R. E. Machol, editors, *Optimal Strategies in Sports*, pages 33–41. North Holland, Amsterdam, 1977.
- R. T. Campbell and D.-S. Chen. A minimum distance basketball scheduling problem. In R. E. Machol, S. P. Ladany, and D. G. Morrison, editors, *Management Science in Sports*, volume 4 of *Studies in the Management Sciences*, pages 15–25. North-Holland, Amsterdam, 1976.
- K. K. H. Cheung. Solving mirrored traveling tournament problem benchmark instances with eight teams. *Discrete Optimization*, 5:138–143, 2008.
- D. Costa. An evolutionary tabu search algorithm and the NHL scheduling problem. *INFOR*, 33:161–178, 1995.
- F. N. Costa, S. Urrutia, and C. C. Ribeiro. An ILS heuristic for the traveling tournament problem with predefined venues. *Annals of Operations Research*, to appear. doi: 10.1007/s10479-010-0719-9.

- D. de Werra. Geography, games, and graphs. *Discrete Applied Mathematics*, 2: 327–337, 1980.
- D. de Werra. Scheduling in sports. In P. Hansen, editor, *Studies on Graphs and Discrete Programming*, pages 381–395. North Holland, Amsterdam, 1981.
- D. de Werra. Minimizing irregularities in sports schedules using graph theory. *Discrete Applied Mathematics*, 4:217–226, 1982.
- D. de Werra. Some models of graphs for scheduling sports competitions. *Discrete Applied Mathematics*, 21:47–65, 1988.
- D. de Werra, J.L. Descombes, and P. Masson. A constrained sports scheduling problem. *Discrete Applied Mathematics*, 26:41–49, 1990.
- F. Della Croce and D. Oliveri. Scheduling the Italian football league: An ILP-based approach. *Computers & Operations Research*, 33:1963–1974, 2006.
- J. H. Dinitz. Starters. In C. J. Colbourn and J. H. Dinitz, editors, *The CRC Handbook of Combinatorial Designs*, The CRC Press Series on Discrete Mathematics and its applications, pages 467–473. CRC Press, Boca Raton, 1996.
- A. R. Duarte and C. C. Ribeiro. Referee assignment in sports leagues: Approximate and exact multi-objective approaches. In *19th International Conference on Multiple Criteria Decision Making*, pages 58–60, Auckland, 2008.
- A. R. Duarte, C. C. Ribeiro, and S. Urrutia. A hybrid ILS heuristic to the referee assignment problem with an embedded MIP strategy. In T. Bartz-Beielstein, M. J. B. Aguilera, C. Blum, B. Naujoks, A. Roli, G. Rudolph, and M. Sampels, editors, *Hybrid Metaheuristics*, volume 4771 of *Lecture Notes in Computer Science*, pages 82–95. Springer, Berlin, 2007a.
- A. R. Duarte, C. C. Ribeiro, S. Urrutia, and E. H. Haeusler. Referee assignment in sports leagues. In E. K. Burke and H. Rudová, editors, *Practice and Theory of Automated Timetabling VI*, volume 3867 of *Lecture Notes in Computer Science*, pages 158–173. Springer, Berlin, 2007b.
- G. Durán, M. Guajardo, J. Miranda, D. Sauré, S. Souyris, A. Weintraub, and R. Wolf. Scheduling the Chilean soccer league by integer programming. *Interfaces*, 37:539–552, 2007a.
- G. Durán, T.F. Noronha, C.C. Ribeiro, S. Souyris, and A. Weintraub. Branch-and-cut for a real-life highly constrained soccer tournament scheduling problem. In E. Burke and H. Rudová, editors, *Practice and Theory of Automated Timetabling VI*, volume 3867 of *Lecture Notes in Computer Science*, pages 174–186. Springer, 2007b.
- G. Durán, M. Guajardo, A. Weintraub, and R. Wolf. O.R. & Soccer: Scheduling the Chilean league using mathematical programming. *OR/MS Today*, 36:42–47, 2009.
- K. Easton, G. Nemhauser, and M. A. Trick. The travelling tournament problem: Description and benchmarks. In T. Walsh, editor, *Principles and Practice of Constraint Programming*, volume 2239 of *Lecture Notes in Computer Science*, pages 580–585. Springer, Berlin, 2001.
- K. Easton, G. L. Nemhauser, and M. A. Trick. Solving the travelling tournament problem: A combined integer programming and constraint programming approach. In E. K. Burke and P. de Causmaecker, editors, *Practice and Theory of Automated Timetabling IV*, volume 2740 of *Lecture Notes in Computer Science*, pages 100–109, Berlin, 2003. Springer.

- J. R. Evans. A microcomputer-based decision support system for scheduling umpires in the American Baseball League. *Interfaces*, 18:42–51, 1988.
- J. R. Evans, J. E. Hebert, and R. F. Deckro. Play ball - The scheduling of sports officials. *Perspectives in Computing*, 4:18–29, 1984.
- A. Farmer, J. S. Smith, and L. T. Miller. Scheduling umpire crews for professional tennis tournaments. *Interfaces*, 37:187–196, 2007.
- J. A. Ferland and C. Fleurent. Computer aided scheduling for a sport league. *INFOR*, 29:14–25, 1991.
- J. Fiallos, J. Pérez, F. Sabillón, and M. Licona. Scheduling soccer league of Honduras using integer programming. In A. Johnson and J. Miller, editors, *Proceedings of the 2010 Industrial Engineering Research Conference*, San Carlos, 2010.
- T. Flatberg, E. Nilssen, and M. Stlevik. Scheduling the topmost fotball leagues of Norway. In *Book of Abstracts of the 23rd European Conference on Operational Research*, page 240, Bonn, 2009.
- C. Fleurent and J. A. Ferland. Allocating games for the NHL using integer programming. *Operations Research*, 41:649–654, 1993.
- D. Fronček. Scheduling the Czech national basketball league. *Congressus Numerantium*, 153:5–24, 2001.
- C. Goodbread. SEC aims to fix schedule problem by month's end, 2010. Online reference at <http://www.tidesports.com/article/20100409/NEWS/100409591>, last visited on August 14, 2010.
- D. Goossens and F. Spieksma. Scheduling the Belgian soccer league. *Interfaces*, 39:109–118, 2009a.
- D. Goossens and F. Spieksma. Does the carry-over effect exist? In *Book of Abstracts of the 23rd European Conference on Operational Research*, page 288, Bonn, 2009b.
- A. Guedes and C. C. Ribeiro. A hybrid heuristic for minimizing weighted carry-over effects in round robin tournaments. In *Proceedings of the 4th Multidisciplinary International Conference on Scheduling Theory and Applications*, pages 115–129, Dublin, 2009.
- M. Henz. Scheduling a major college basketball conference - Revisited. *Operations Research*, 49:163–168, 2001.
- M. Henz, T. Müller, and S. Thiel. Global constraints for round robin tournament scheduling. *European Journal of Operational Research*, 153:92–101, 2004.
- R. Hoshino and K. Kawarabayashi. The multi-round balanced traveling tournament problem. In *21st International Conference on Automated Planning and Scheduling*, Freiburg, 2011a. To appear.
- R. Hoshino and K. Kawarabayashi. The inter-league extension of the traveling tournament problem and its application to sports scheduling. In *25th AAAI Conference on Artificial Intelligence*, San Francisco, 2011b. To appear.
- R. Hoshino and K. Kawarabayashi. The distance-optimal inter-league schedule for Japanese pro baseball. In *Workshop on Constraint Satisfaction Techniques for Planning and Scheduling Problem*, 21st International Conference on Automated Planning and Scheduling, Freiburg, 2011c. To appear.
- S. Irnich. A new branch-and-price algorithm for the traveling tournament problem. *European Journal of Operational Research*, 204:218–228, 2010.
- G. Kendall, S. Knust, C.C. Ribeiro, and S. Urrutia. Scheduling in sports: An annotated bibliography. *Computers and Operations Research*, 37:1–19, 2010.

- T. Kirkman. On a problem in combinations. *Cambridge Dublin Math Journal*, 2: 191–204, 1847.
- S. Knust. Scheduling non-professional table-tennis leagues. Technical Report 270, Osnabrücker Schriften zur Mathematik, 2007.
- S. Knust. Scheduling non-professional table-tennis leagues. *European Journal of Operational Research*, 200:358–367, 2010.
- A. Lamghari and J. A Ferland. In D. Srinivasan, editor, *Proceedings of IEEE Symposium on Computational Intelligence in Scheduling*, pages 238–245, Honolulu, 2007.
- A. Lamghari and J. A Ferland. *European Journal of Operational Research*, 210: 694–705, 2010a.
- A. Lamghari and J. A Ferland. *Annals of Operational Research*, 180:33–61, 2010b.
- R. A. Melo, S. Urrutia, and C. C. Ribeiro. The traveling tournament problem with predefined venues. *Journal of Scheduling*, 12:607–622, 2009.
- R. Miyashiro and T. Matsui. A polynomial-time algorithm to find an equitable home-away assignment. *Operations Research Letters*, 33:235–241, 2005.
- R. Miyashiro and T. Matsui. Minimizing the carry-over effects value in a round robin tournament. In *Proceedings of the 6th International Conference on the Practice and Theory of Automated Timetabling*, pages 460–463, Brno, 2006.
- R. Miyashiro, H. Iwasaki, and T. Matsui. Characterizing feasible pattern sets with a minimum number of breaks. In E. K. Burke and P. de Causmaecker, editors, *Practice and Theory of Automated Timetabling IV*, volume 2740 of *Lecture Notes in Computer Science*, pages 78–99, Berlin, 2003. Springer.
- G. L. Nemhauser and M. A. Trick. Scheduling a major college basketball conference. *Operations Research*, 46:1–8, 1998.
- K. Nurmi, D. Goossens, T. Bartsch, F. Bonomo, D. Briskorn, G. Duran, J. Kyngäs, J. Marenco, C. C. Ribeiro, F. Spieksma, S. Urrutia, and R. Wolf. A framework for a highly constrained sports scheduling problem. In *Proceedings of the International Multi-Conference of Engineers and Computer Scientists*, volume III, pages 1991–1997, Hong-Kong, 2010.
- R. Ordonez and T. W. Knowles. Solving the American League umpire crew scheduling problem as a constraint satisfaction problem. In *Proceedings of the 29th Annual Meeting of the Decision Sciences Institute*, volume 2, pages 1058–1061, Las Vegas, 1998.
- G. Post and G. J. Woeginger. Sports tournaments, home-away assignments, and the break minimization problem. *Discrete Optimization*, 3:165–173, 2006.
- R. V. Rasmussen. Scheduling a triple round robin tournament for the best Danish soccer league. *European Journal of Operational Research*, 185:795–810, 2008.
- R. V. Rasmussen and M. A. Trick. The timetable constrained distance minimization problem. In J. C. Beck and Smith B. M., editors, *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, volume 3990 of *Lecture Notes in Computer Science*, pages 167–181. Springer, 2006.
- R. V. Rasmussen and M. A. Trick. A Benders approach for the constrained minimum break problem. *European Journal of Operational Research*, 177:198–213, 2007.
- R. V. Rasmussen and M. A. Trick. Round robin scheduling – A survey. *European Journal of Operational Research*, 188:617–636, 2008.

- J.-C. Regin. Minimization of the number of breaks in sports scheduling problems using constraint programming. In E. Freuder and R. Wallace, editors, *Constraint Programming Large Scale Discrete Optimization*, volume 57 of *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 115–130, 2001.
- C. C. Ribeiro and S. Urrutia. Scheduling the Brazilian soccer tournament with fairness and broadcast objectives. In E. K. Burke and H. Rudová, editors, *Practice and Theory of Automated Timetabling VI*, volume 3867 of *Lecture Notes in Computer Science*, pages 147–157. Springer, Berlin, 2007a.
- C. C. Ribeiro and S. Urrutia. Heuristics for the mirrored traveling tournament problem. *European Journal of Operational Research*, 179:775–787, 2007b.
- C. C. Ribeiro and S. Urrutia. Bicriteria integer programming approach for scheduling the Brazilian national soccer tournament. In *Proceedings of the Third International Conference on Management Science and Engineering Management*, pages 46–49, Bangkok, 2009.
- C. C. Ribeiro and S. Urrutia. Soccer scheduling goaaaaal! *OR/MS Today*, 37(2): 54–59, 2010.
- C. C. Ribeiro and S. Urrutia. Scheduling the Brazilian soccer tournament: Solution approach and practice. *Interfaces*, 2011. To appear.
- K. G. Russell. Balancing carry-over effects in round robin tournaments. *Biometrika*, 67:127–131, 1980.
- R. A. Russell and J. M. Leung. Devising a cost effective schedule for a baseball league. *Operations Research*, 42:614–625, 1994.
- J. Schönberger, D. C. Mattfeld, and H. Kopfer. Memetic algorithm timetabling for non-commercial sport leagues. *European Journal of Operational Research*, 153: 102–116, 2004.
- J. A. M. Schreuder. Combinatorial aspects of construction of competition Dutch professional football leagues. *Discrete Applied Mathematics*, 35:301–312, 1992.
- J.A.M. Schreuder. Constructing timetables for sport competitions. *Mathematical Programming Study*, 13:58–67, 1980.
- C. Thielen and S. Westphal. Complexity of the traveling tournament problem. *Theoretical Computer Science*, 412:345–351, 2011.
- M. A. Trick. A schedule-then-break approach to sports timetabling. In E. K. Burke and W. Erben, editors, *Selected papers from the Third International Conference on Practice and Theory of Automated Timetabling III*, volume 2079 of *Lecture Notes in Computer Sciences*, pages 242–253. Springer, 2000.
- M. A. Trick. Integer and constraint programming approaches for round robin tournament scheduling. In E. Burke and P. de Causmaecker, editors, *Practice and Theory of Automated Timetabling IV*, volume 2740 of *Lecture Notes in Computer Science*, pages 63–77, Berlin, 2003. Springer.
- M. A. Trick. Formulations and reformulations in integer programming. In R. Barták and M. Milano, editors, *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, volume 3524 of *Lecture Notes in Computer Science*, pages 833–836. Springer, 2005.
- M. A. Trick. Challenge traveling tournament instances, 2010. Online reference at <http://mat.tepper.cmu.edu/TOURN/>, last visited on May 19, 2010.
- S. Urrutia and C. C. Ribeiro. Maximizing breaks and bounding solutions to the mirrored traveling tournament problem. *Discrete Applied Mathematics*, 154:1932–1938, 2006.

- S. Urrutia, C. C. Ribeiro, and R. A. Melo. A new lower bound to the traveling tournament problem. In *Proceedings of the IEEE Symposium on Computational Intelligence in Scheduling*, pages 15–18, Honolulu, 2007. IEEE.
- D. C. Uthus, P. J. Riddle, and H. W. Guesgen. Solving the traveling tournament problem with iterative-deepening A*. *Journal of Scheduling*, 2011. To appear.
- P. Van Hentenryck and Y. Vergados. Traveling tournament scheduling: A systematic evaluation of simulated annealing. In *Integration of AI and OR Techniques in Constraint Programming for Combinatorial Optimization Problems*, volume 3990 of *Lecture Notes in Computer Science*, pages 228–243. Springer, Berlin, 2006.
- R. J. Willis and B. J. Terrill. Scheduling the Australian state cricket season using simulated annealing. *Journal of the Operational Research Society*, 45:276–280, 1994.
- M. Wright. Scheduling English cricket umpires. *Journal of the Operational Research Society*, 42:447–452, 1991.
- M. Wright. A fair allocation of county cricket opponents. *Journal of the Operational Research Society*, 43:195–201, 1992.
- M. Wright. Scheduling fixtures for New Zealand cricket. *IMA Journal of Management Mathematics*, 16:99–112, 2005.
- M. Wright. Scheduling fixtures for basketball New Zealand. *Computers & Operations Research*, 33:1875–1893, 2006.
- M. Yavuz, U. H. Inan, and A. Figlali. Fair referee assignments for professional football leagues. *Computers & Operations Research*, 35:2937–2951, 2008.

(Celso C. Ribeiro) DEPARTMENT OF COMPUTER SCIENCE, UNIVERSIDADE FEDERAL FLUMINENSE,
RUA PASSO DA PÁTRIA 156, NITERÓI, RJ 24210-240 BRAZIL.

E-mail address: celso@ic.uff.br