Scheduling the Brazilian soccer tournament with fairness and broadcast objectives

Celso C. Ribeiro[1] and Sebastián Urrutia[2]

[1] Department of Computer Science, Universidade Federal Fluminense, Rua Passo da Pátria 156, Niterói, RJ 24210-240, Brazil.
E-mail: celso@inf.puc-rio.br

[2] Department of Computer Science, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, Belo Horizonte, MG 31270-010, Brazil.
E-mail: surrutia@dcc.ufmg.br

Abstract. The Brazilian soccer tournament is organized every year by the Brazilian Soccer Confederation. Its major sponsor is TV Globo, the largest media group and television network in Brazil, who imposes constraints on the games to be broadcast. Scheduling the games of this tournament is a very constrained problem, with two objectives: breaks minimization (fairness) and the maximization of the revenues from TV broadcasting. We propose an integer programming decomposition strategy to solve this problem to optimality. Numerical results obtained for the 2005 and 2006 editions of the tournament are reported and compared.

1 Introduction

Soccer is the most widely practiced sport in Brazil. The yearly Brazilian soccer tournament is the most important sport event in the country. It is organized by the Brazilian Soccer Confederation (CBF). Its major sponsor is TV Globo, the largest media group and television network in Brazil, who imposes constraints on the games to be broadcast.

The most attractive games are those involving a subset of elite teams with more fans and, consequently, with larger broadcast shares. Games involving teams from São Paulo and Rio de Janeiro (the two largest cities in Brazil) are of special interest to TV Globo, due to larger revenues from advertising.

The competition lasts seven months and is structured as a compact mirrored double round robin (MDRR) tournament [3]. It is played by n teams, where n is an even number ($n = 24$ in 2004, $n = 22$ in 2005, and $n = 20$ in 2006). There are $2n - 2$ rounds and each team plays exactly once in each round. There are at most two rounds of games per week. Each team faces every other twice: once at home and the other away. If team a plays against team b at home (resp. away) in round k, with

$k < n$, then team a plays against team b away (resp. at home) in round $k + n - 1$. We refer to [3] for a recent survey on the sport scheduling literature.

The revenues and the attractiveness of the tournament strongly depend on the schedule of the games. The organizers and the sponsors search for a schedule optimizing two different objectives. CBF attempts to maximize fairness, by minimizing the number of breaks along the tournament (breaks minimization objective). A break occurs whenever a teams plays two consecutive home games or two consecutive away games, see e.g. [8]. TV Globo aims to maximize its revenues, by maximizing the number of relevant games it is able to broadcast (broadcast objective). The schedule must also satisfy a number of hard constraints.

We propose an integer programming solution approach for solving this scheduling problem, based on the generation of feasible home-away patterns. The detailed problem formulation is presented in Section 2. The solution strategy is described in Section 3. Numerical results obtained for real-life instances corresponding to the 2005 and 2006 editions of the tournament are reported and compared in Section 4. Concluding remarks are drawn in the last section.

2 Problem statement

We consider both the 2005 and 2006 editions of the competition, with respectively $n = 22$ and $n = 20$ participating teams. Every team has a home city and some cities host more than one team. Some teams are considered and handled as elite teams, due to their number of fans, to the records of their previous participations in the tournament, and to the value of their players. There are weekend rounds and mid-week rounds.

São Paulo and Rio de Janeiro are the two largest cities in Brazil (with more fans and, consequently, generating larger revenues from advertising) and both of them have four elite teams. Games cannot be broadcast to the same city where they take place and only one game per round can be broadcast to each city. Consequently, TV Globo wants to broadcast to São Paulo (resp. Rio de Janeiro) games in which an elite team from São Paulo (resp. Rio de Janeiro) plays away against another elite team from another city. Such games will be referred to as TV games.

Belém is a city very far away from São Paulo and Rio de Janeiro. TV Globo is not willing to broadcast games taking place at Belém, due to the high logistical costs. Besides following the structure of a MDRR tournament, the schedule should also satisfy other hard constraints:

1. Every team playing at home (resp. away) in the first round plays away (resp. at home) in the last round.
2. Every team plays once at home and once away in the two first rounds and in the two last rounds.

3. After any number of rounds during the first half of the tournament, the difference between the number of home games and away games played by any team is either zero or one (i.e., the number of home and away games is always balanced in the first $n-1$ rounds).
4. Some pairs of teams with the same home city have complementary patterns (i.e., whenever one of them plays at home, the other plays away).
5. Flamengo and Fluminense (two elite teams from Rio de Janeiro that share the same stadium for their home games) have complementary patterns in the last four rounds.
6. Games between teams from the same city are not to be played in mid-week rounds or in the last six rounds (since they are among the most attractive games).
7. There is at least one elite team from Rio de Janeiro playing outside Rio de Janeiro and one elite team from São Paulo playing outside São Paulo in every round.
8. If in some round there is only one elite team from Rio de Janeiro (resp. São Paulo) playing outside Rio de Janeiro (resp. São Paulo), then this game should not be held in Belém.

The two objectives that must be optimized are the minimization of the number of breaks and the maximization of the number of rounds in which there is at least one TV game to be broadcast to São Paulo plus the number of rounds in which there is at least one TV game to be broadcast to Rio de Janeiro. Therefore, while the broadcast objective regards the number of relevant games that TV Globo is able to broadcast, the breaks minimization objective establishes the home-away equilibrium in the sequence of games played by each team.

The requirements described above in terms of constraints and objectives resulted from several meetings and discussions with the organizers of the tournament and, in particular, with officials of TV Globo.

3 Solution strategy

We propose the following approach to tackle this bi-objective problem. First, we add to the problem an extra constraint stating that the number of breaks is fixed at its minimum. Then, the broadcast objective is maximized with this additional constraint. If the maximum objective value of this restricted problem is equal to the unconstrained maximum, then this solution is what in multi-criteria optimization [4] is called an ideal point (all objectives are at their individual optimal values simultaneously). If the maximum solution value of the restricted problem is smaller than the unconstrained maximum, then the solution of the first is not an ideal point but still is a non-dominated solution (no other solution is better with respect to one of the objectives without being worse with respect to the other). Figure 1 illustrates this approach.

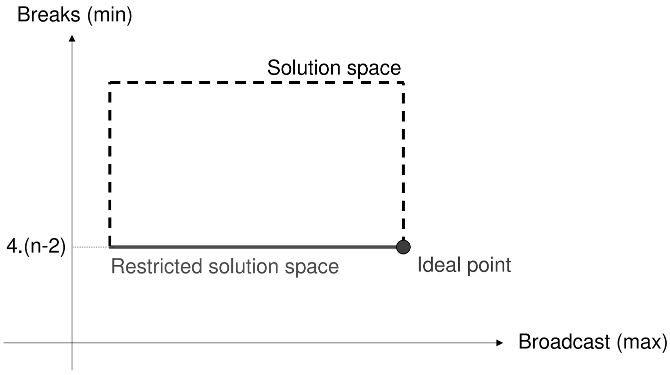

Fig. 1. Solution space and restricted solution space.

We assume that there exists at least one feasible solution with a minimum number of breaks, i.e., we assume that the restricted search space illustrated in Figure 1 is not empty. This fact was experimentally verified for all test instances and their variations.

3.1 Bounds

Given a round robin schedule with r rounds, a home-away pattern (HAP) is a vector of r positions filled with "A"s and "H"s. An "A" (resp. "H") in position s of a HAP indicates that every team associated with this HAP plays away (resp. at home) at round s. Since each team has to play against every other team, each team must be associated with a different HAP. Figure 2 shows a HAP set for six teams in a single round robin schedule.

Team 1:	A H A H A
Team 2:	H A H A H
Team 3:	A H H H A
Team 4:	H A A H H
Team 5:	H A A A H
Team 6:	A H H A A

Fig. 2. HAP set for a tournament with six teams.

We first show that constraints (1) and the mirrored structure impose that $4(n-2)$ is a lower bound to the number of breaks.

Since the tournament is mirrored, if the number of breaks in the first half of a HAP is even, then the total number of breaks is also even and equal to twice the number of breaks in the first half. On the other hand, if the number of breaks in the first half is odd, then the total number of breaks is also odd and equal to twice the number of breaks in the first half plus one (there is an extra break in the first round of the second half).

There are only two HAPs without breaks for single round robin schedules, one starting with a home game and the other starting with an away game. Therefore, the other $n-2$ teams must have at least one break in the first half of the schedule, yielding at least three breaks in the whole schedule. We notice that in double round robin schedules a team with an odd number of breaks plays its last game in the same playing condition of its first game (home-home or away-away). Therefore, to satisfy constraints (1), we shall consider schedules in which every team has an even number of breaks. In consequence, the $n-2$ teams having three breaks must have an extra break and the number of breaks cannot be smaller than $4(n-2)$.

The broadcast objective can also be bounded. Since at most one TV game can be broadcast to São Paulo and another to Rio de Janeiro in every round, the broadcast objective cannot be greater than twice the number of rounds. Furthermore, the broadcast objective is also bounded by the number of existing TV games. The later is equal to the number of elite teams from São Paulo multiplied by the number of elite teams outside São Paulo, plus the number of elite teams from Rio the Janeiro multiplied by the number of elite teams outside Rio de Janeiro. The second bound is stronger (i.e., smaller) for the instances solved in this work. As an example, the first bound is equal to 84 for $n = 22$, while the second is equal to 56 (four elite teams from Rio de Janeiro, four from São Paulo, and three from other cities).

3.2 Solution algorithm

An straightforward integer programming formulation of the problem could not be solved by a commercial solver such as CPLEX after an entire day of computations.

Decomposition methods have been previously proposed for problems where the distances between the venues were not relevant. Nemhauser and Trick [6] proposed a three-phase scheme to exactly solve the problem of scheduling a basketball league. Feasible home-away patterns are created in the first phase. In the second phase, a different feasible HAP is assigned to each team (two different teams must have different HAPs in every feasible round robin schedule). Finally, in the last phase, the schedule is created respecting the previously determined HAP assignments.

Some recent papers dealt with scheduling problems in soccer tournaments. Della Croce and Oliveri [2] tackled the Italian soccer league. Bartsch et al. [1] worked on the schedule of the soccer leagues of Austria and Germany. Goossens and Spieksma [5] considered the scheduling of the Belgian soccer league. Noronha et al. [7] proposed a branch-and-cut algorithm to schedule the Chilean soccer tournament. The algorithms proposed in [1, 2] follow a decomposition scheme similar to that of [6].

We propose an algorithm following an approach similar to the above described multi-phase decomposition scheme. Figure 3 illustrates this approach, whose four phases are described in the next sections.

3.3 Phase 1: HAP generation

HAPs of mirrored schedules are divided into two symmetric halves. The second half is completely determined by the first. Therefore, we may determine which properties the first half of a HAP must obey so as that the entire HAP be feasible. As noticed in Section 3.1, if the number of breaks in the first half of a HAP is even, then the total number of breaks is also even and equal to twice the number of breaks in the first half. On the other hand, if the number of breaks in the first half is odd, then the total number of breaks is also odd and equal to twice the number of breaks in the first half plus one.

HAPs satisfying constraints (1) are those with an even number of breaks. Therefore, we consider only HAPs with an even number of breaks in the first half. Since we are interested in schedules with a minimum number of breaks, we only consider HAPs with either zero (there are only two such HAPs) or two breaks in the first half.

HAPs satisfying constraints (2) are those without breaks in the second and last rounds. Since the schedule is mirrored, they must have no breaks in the last round of the first half (round $n - 1$).

The difference between the number of home and away games in a HAP without breaks is equal to one after odd rounds and equal to zero after even rounds. If a HAP has a break in an even round, this difference increases to two. Therefore, HAPs satisfying constraints (3) are those without breaks in even rounds of the first half. Even (resp. odd) rounds of the second half are globally odd (resp. even) in mirrored schedules. In consequence, teams with breaks have at least one break in an even round and the difference between the number of home and away games will be necessarily greater than one after some round. For this reason, constraints (3) are limited to the first half of the schedule. If they were imposed to the whole schedule, the problem would be infeasible.

Consequently, feasible HAPs for the first half are those without breaks or with exactly two breaks in odd rounds (but not in the last, since constraints (2) forbid breaks in the last round of the second half). There are $n/2 - 2$ rounds (all odd rounds but the first and the last) in which teams may have their two breaks, yielding a total of $\binom{n/2-2}{2} = (n/2-2) \cdot (n/2-3)/2$ possible break configurations. Since there are two HAPs for

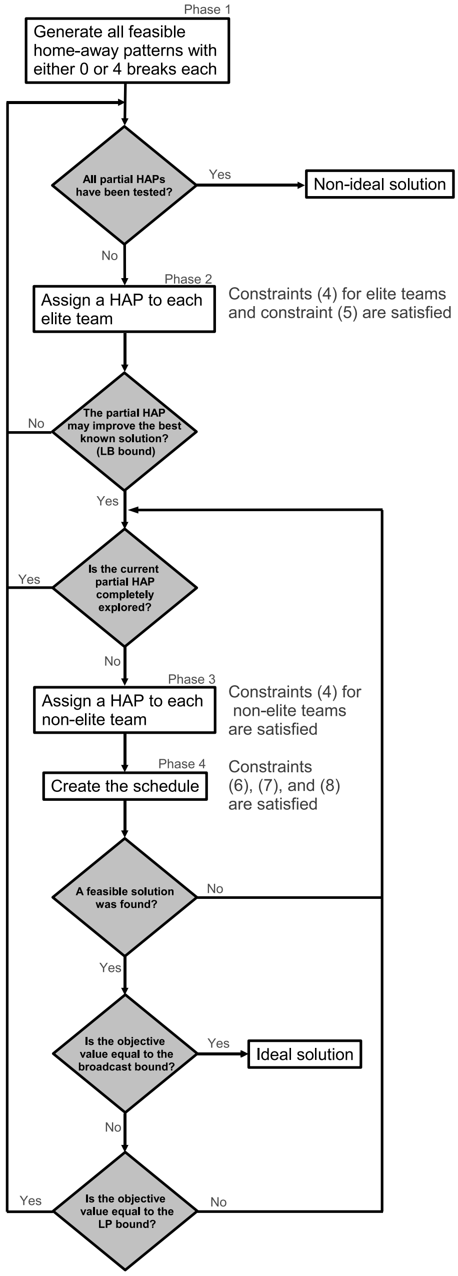

Fig. 3. Solution approach.

every possible break configuration (one starting by a home game and the other by an away game), the number of feasible HAPs with two breaks is equal to $(n/2-2)\cdot(n/2-3)$. Considering the two HAPs without breaks, the total number of feasible HAPs is equal to $(n/2-2)\cdot(n/2-3)+2$.

The number of feasible HAPs is equal to 58 for $n = 20$ and to 74 for $n = 22$. This small number of feasible home-away patterns with at most two breaks each allows their complete enumeration in this phase.

3.4 Phase 2: assignment of partial HAPs to elite teams

In this phase, we use an explicit exhaustive enumeration to assign a HAP to each elite team satisfying constraints (4) and (5). The use of the two HAPs with no breaks is enforced, to keep the number of breaks at its minimum value $4(n-2)$.

Constraint (4) is satisfied by assigning complementary HAPs to every pair of teams to which this constraint is imposed. Since half of the teams play at home and half away in every round of a feasible HAP assignment, in this phase and the next we first enumerate HAP assignments in which if one pattern is used, the complementary pattern is used as well. In this way, we improve the chance of a HAP assignment to be feasible.

Since Flamengo and Fluminense have their own complementary teams, they cannot have complementary patterns themselves. Therefore, to satisfy constraint (5) we assign patterns in such a way that (i) if Flamengo starts at home (resp. away), then Fluminense starts away (resp. at home), and (ii) either both or none of them have a break in round $n-3$. This ensures that Flamengo and Fluminense will have complementary patterns in the last four rounds, therefore satisfying constraint (5).

After assigning a HAP to each elite team, we build and solve a linear programming model considering the partial HAP assignments already made (we remind that they satisfy constraints (1) to (5)), maximizing the broadcast objective subject to constraints (6) to (8). This linear programming model enforces that each elite team will play either at home or away in each round, depending on the HAP assigned to it.

The optimal value of the above linear program is an upper bound to the broadcast objective, associated to the current partial HAP assignment. This current partial assignment can be discarded if the bound it provides is not smaller than the value of the broadcast objective for the current best known feasible solution. In this case, a new partial HAP assignment is enumerated and this phase is repeated for the new assignment. Otherwise, the algorithm proceeds to the third phase.

If all partial HAP assignments have been enumerated, the algorithm stops and returns the best feasible solution it obtained. The latter is non-dominated, but not an ideal solution.

3.5 Phase 3: assignment of HAPs to non-elite teams

In this phase, once again we use an explicit exhaustive enumeration to assign one of the still available HAPs to each non-elite team, satisfying constraint (4) and completing the partial assignment constructed in the previous phase. Once the HAP assignment is complete, the algorithm proceeds to the last phase.

Whenever all possible alternatives to complete the partial assignment of HAPs to elite teams have been tested, the algorithm goes back to the second phase to enumerate a new partial assignment.

3.6 Phase 4: schedule creation

At this point, there is a HAP assigned to each team. In this phase, we build and solve an integer programming problem considering the current HAP assignments, maximizing the broadcast objective subject to constraints (6) to (8). This integer program defines the games to be played in each round, according to the home-away patterns assigned to each team.

In this straightforward problem formulation, we use binary variables x_{ijk} that are equal to one if and only if team i plays with team j at home in round k. Since at this phase we already know which teams play at home and which play away in each round, half of the variables are trivially equal to zero and can be eliminated. We also know that, due to the home-away pattern assigned to each team, the number of breaks is minimum and equal to $4(n-2)$. Therefore, the number of variables is relatively small, since there is no need to use further variables to model the breaks. In consequence, this model can be quickly solved by a commercial solver.

We first assume that the above integer program is feasible. If its optimal value is equal to the broadcast bound, then the algorithm terminates with an ideal solution. Otherwise, if the optimal value of the integer program is equal to the linear programming bound obtained in the second phase, then the algorithm returns to the second phase to enumerate a new partial HAP assignment, because no better solution can be obtained with the current partial assignment.

If the integer program is infeasible or if its optimal value is smaller than the linear programming bound, then the algorithm returns to the third phase to enumerate another complete HAP assignment.

In any case, if the integer program is feasible and its optimal value is smaller than that of the current best known feasible solution, then the latter is updated.

4 Application to real-life instances

The algorithm was applied to two instances corresponding to the 2005 and 2006 editions of the Brazilian tournament, with 22 and 20 teams respectively. There were 11 elite teams in each instance: four from São Paulo, four from Rio de Janeiro, and three from other cities.

CPLEX 9.0 was used as the linear and integer programming solver in the computational experiments. The algorithm was coded in C++, compiled with gcc and executed on a standard Pentium IV processor with 256 Mbytes of RAM memory.

The approach proposed in this work was able to compute optimal schedules providing ideal solutions (i.e., simultaneously optimizing both the broadcast and the breaks objectives) for the 2005 and 2006 editions of the tournament, always in less than ten minutes of execution time.

Tables 1 and 2 compare the schedules produced by the new algorithm (HAP-ILP) with those elaborated by the tournament organizers using ad hoc procedures based on their own expertise, acquired after many years creating the tournament schedule. The schedules are compared in terms of the satisfaction of the problem constraints and of the values of the two objective functions.

The solutions produced by our four-phase approach are clearly better than those produced by the current scheduler. Three types of constraints were not fully satisfied in the official schedules of the 2005 and 2006 editions of the tournament, while in both cases our algorithm provided optimal solutions satisfying all constraints. Regarding the 2005 edition, the ad hoc rules lead to schedules with 152 breaks and 43 TV games to be broadcast. The four-phase integer programming approach proposed in this work found a schedule with only 80 breaks (which is optimal), in which all 56 TV games could be broadcast (which is once again optimal). A similar comparison can be done for the 2006 edition of the tournament.

5 Conclusions

In this paper, we proposed a four-phase integer programming approach for scheduling the yearly Brazilian soccer tournament. This is a bi-criteria highly constrained mirrored round robin tournament, combining a fairness objective established by the organizers with an economical objective imposed by the TV sponsors.

The proposed algorithm was able to obtain optimal solutions maximizing both objectives and satisfying all constraints in a few minutes of computation time on a standard desktop computer. The schedules provided by the four-phase approach are clearly better than those currently used by the tournament organizers, which are obtained by simple ad

	Official schedule	HAP-ILP schedule
Constraints (1)	yes	yes
Constraints (2)	yes	yes
Constraints (3)	no	yes
Constraints (4)	yes	yes
Constraint (5)	no	yes
Constraints (6)	no	yes
Constraints (7)	yes	yes
Constraints (8)	yes	yes
Breaks	156	80 (optimal)
Broadcast	43	56 (optimal)

Table 1. 2005 edition of the Brazilian tournament (22 teams).

	Official schedule	HAP-ILP schedule
Constraints (1)	yes	yes
Constraints (2)	yes	yes
Constraints (3)	no	yes
Constraints (4)	yes	yes
Constraint (5)	no	yes
Constraints (6)	no	yes
Constraints (7)	yes	yes
Constraints (8)	yes	yes
Breaks	172	72 (optimal)
Broadcast	47	56 (optimal)

Table 2. 2006 edition of the Brazilian tournament (20 teams).

hoc rules that are not even able to found feasible solutions satisfying all constraints.

A software system implementing a simple decision support system based on the proposed algorithm is able to generate a collection of same cost solutions to be evaluated and compared by the user. The use of this decision support system and the schedules created with the approach proposed in this work are currently under consideration by the tournament organizers.

References

1. T. Bartsch, A. Drexl, and S. Kröger. Scheduling the professional soccer leagues of Austria and Germany. *Computers and Operations Research*, 33:1963–1974, 2006.

2. F. Della Croce and D. Oliveri. Scheduling the Italian football league: An ILP-based approach. *Computers and Operations Research*, 33:1963–1974, 2006.
3. K. Easton, G.L. Nemhauser, and M. Trick. Sports scheduling. In J.T. Leung, editor, *Handbook of Scheduling: Algorithms, Models and Performance Analysis*, pages 52.1–52.19. CRC Press, 2004.
4. M. Ehrgott. *Multiple Criteria Optimization: Classification and Methodology*. Shaker, 1997.
5. D. Goossens and F. Spieksma. Schedulinfg the Belgian soccer league. In *Proceedings of the 6th International Conference on the Practice and Theory of Automated Timetabling*, pages 420–422, Brno, 2006.
6. G.L. Nemhauser and M.A. Trick. Scheduling a major college basketball conference. *Operations Research*, 46:1–8, 1998.
7. T.F. Noronha, C.C. Ribeiro, G. Duran, S. Souyris, and A. Weintraub. Branch-and-cut for a real-life highly constrained soccer tournament scheduling problem. In *Proceedings of the 6th International Conference on the Practice and Theory of Automated Timetabling*, pages 398–401, Brno, 2006.
8. S. Urrutia and C.C. Ribeiro. Maximizing breaks and bounding solutions to the mirrored traveling tournament problem. *Discrete Applied Mathematics*, 154:1932–1938, 2006.