ISBN 9781846260032, England, UK
Proceedings of The Third International Conference on Management Science and Engineering Management (pp. 46-49)

Bicriteria integer programming approach for scheduling the Brazilian national soccer tournament*

Celso C. Ribeiro[1†], Sebastián Urrutia[2]

[1] Universidade Federal Fluminense, 22410-240, Brazil
[2] Universidade Federal de Minas Gerais, 31270-010, Brazil

ICMSEM Inpress

Abstract — The yearly Brazilian soccer tournament is a compact mirrored double round robin tournament played by 20 teams. Its fixture should satisfy more than 40 different types of hard and soft constraints, ranging from fairness to security issues, and from technical to broadcasting criteria. Soft constraints are handled through penalties and incentives. The main optimization criterion consists in maximizing the number of games played by elite teams that can be broadcast by open TV channels, so as to maximize broadcast audience and broadcast rights. A decomposition strategy is used to solve the associated integer programming formulation. The resulting automatic system was tuned with data from the 2005 and 2006 editions of the tournament and fully used in practice for the first time to schedule the 2009 tournament, providing several alternative choices that were offered for the user's final decision.

Keywords: scheduling, round robin, integer programming, decomposition

1 Motivation and background

Soccer is the most widely practiced sport in Brazil. The Brazilian national soccer tournament organized every year by the Brazilian Soccer 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. The most attractive games are those involving teams with more fans and better players, and, consequently, also with larger broadcast shares. Games involving teams from São Paulo and Rio de Janeiro are of special interest for broadcasting through open TV channels, due to their corresponding larger revenues from advertising.

The competition lasts for seven months (from May to December) and is structured as a compact mirrored double round robin tournament [4] played by $n = 20$ teams. Every team has a home city and some cities host more than one team. Each team faces every other twice: once at home and the other away. If team i plays against team j at home (resp. away) in round k, with $1 \leq k \leq n-1$, then team i plays against team j away (resp. at home) in round $k+n-1$. Therefore, the tournament is divided into two phases: any team $i = 1, \cdots, n$ faces every other team $j \neq i$ once in the first phase (rounds 1 to $n-1$) and once in the second phase (rounds n to $2n-2$). There are at most two rounds of games per week: mid-week rounds are played on Wednesdays and Thursdays, while weekend rounds are played on Saturdays and Sundays. Elite teams are those with larger numbers of fans, better records of previous participations in the tournament, and more valuable players. The most important games involve elite teams and should as much as possible be played during the weekends, when they can attract larger attendance and TV audience. Participating teams and the dates available for game playing change from every year to the next. We refer to [5, 9] for recent surveys on sports scheduling.

Revenues from broadcast rights, public attendance, and tournament attractiveness strongly depend on the schedule of the games, which must also satisfy a number of hard and soft constraints. The organizers and the sponsors search for a schedule optimizing two different objectives:

- fairness maximization, by minimizing the total number of consecutive home games and consecutive away games played by any team along the tournament, and
- revenue maximization, by maximizing the number of relevant games that open TV channels are able to broadcast.

We propose an integer programming solution approach for solving this scheduling problem, based on the generation of feasible home-away patterns of game playing. Optimization techniques such as integer programming and heuristics have been applied to the scheduling of soccer tournaments in other contexts, see e.g. [1–3, 7, 8]. The main requirements that should be met by the fixtures of the Brazilian national tournament are reported in Section 2, as determined by the Brazilian Soccer Confederation. The basic structure of the integer programming model is summarized in Section 3. A three-phase solution optimization approach is described in Section 4. Numerical results obtained with real-life data corresponding to the 2005 and

* We are grateful to one anonymous referee whose remarks contributed to improve the original manuscript. The mathematical model and the software system described in this work have been developed in partnership with the Brazilian Soccer Confederation (CBF). All rights belong to the authors. Research of the first author was sponsored by grant 301694/2007-9 of the Brazilian National Council for Scientific and Technological Development (CNPq) and by FAPERJ grant E-26/102.805/2008. Work of the second author was supported by CNPq research grant 302560/2007-6.

† *Correspondence to*: *E-mail address*: celso@ic.uff.br.

2006 editions of the tournament are reported in Section 5. Finally, the last section presents a short account of the current use in practice of the proposed approach. We also give some concluding remarks with respect to the use of exact methods and heuristics in the solution of scheduling problems in sports.

2 Requirements and optimization criteria

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, sponsors want to be able to broadcast to São Paulo (resp. Rio de Janeiro) the games in which an elite team from São Paulo (resp. Rio de Janeiro) plays away against another elite team from another city.

Some of the constraints are summarized below. Most of them reflect strategies for maximizing revenues and attractiveness, while others attempt to avoid unfair situations that could privilege one team or another with a more convenient schedule. These requirements have been discussed and established by teams, federations, city administrators, security forces, and sponsors over the years:

(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 first two rounds and in the last two rounds.

(3) After any number of rounds during the second half of the tournament, the difference between the number of home games and away games to be played by any team is either zero or one.

(4) 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) Games between elite teams in the same city are not to be played in mid-week rounds, in the first two rounds, or in the last four rounds.

(6) No team can play more than two consecutive home games or more than two consecutive away games.

(7) No team should play more than five consecutive games against elite teams.

(8) 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.

(9) Several other constraints handle specific or regional requirements, privileged dates, etc.

The first objective to be optimized is the minimization of breaks, i.e. the total number of consecutive home games and consecutive away games played by any team. It corresponds to seeking a fixture in which the teams alternate home games with away games as much as possible. The second objective is 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. It leads to the maximization of TV audience and revenues from advertising. Soft constraints are handled via penalties and incentives incorporated into the objective function as a weighted sum, with their respective costs and profits significantly smaller than those associated with the two main objectives.

3 Problem modeling

Integer programming is a useful tool to model and solve sports scheduling problems. Some round robin tournament scheduling problems can be solved by directly applying an integer programming solver to the model. In most cases, however, decomposition schemes are used to tackle each stage of the problem by integer programming or other techniques such as constraint programming, complete enumeration, or heuristics.

Given that n denotes the number of teams, there are exactly $r = 2n - 2$ rounds in a compact mirrored double round robin tournament. We define the binary variables

$$x_{ijt} = \begin{cases} 1, & \text{if team } i \text{ plays at home against team } j \text{ in round } t, \\ 0, & \text{otherwise,} \end{cases}$$

for teams $i, j = 1, \cdots, n$ (with $i \neq j$) and rounds $t = 1, \cdots, r = 2n - 2$. The structural constraints of the compact mirrored double round robin tournament may then be formulated as

$$\sum_{t=1}^{2n-2} x_{ijt} = 1, \qquad \forall\, 1 \le i, j \le n, i \neq j \tag{1}$$

$$\sum_{t=1}^{n-1} (x_{ijt} + x_{jit}) = 1, \qquad \forall\, 1 \le i, j \le n, i \neq j \tag{2}$$

$$\sum_{j=1}^{n} x_{ijt} = 1, \qquad \forall\, 1 \le i \le n, 1 \le t \le 2n - 2. \tag{3}$$

The first set of constraints imposes that every team must play against every other team at home exactly once. Constraints in the second set determine that each pair of teams face each other once in the first phase (rounds 1 to $n - 1$) and once in the second phase (rounds n to $2n - 2$). The third set establishes that every team plays exactly once in each round.

The problem has more than 40 different types of practical constraints, as summarized in Section 2. Besides the structural constraints (1) to (3), all other constraints handle specific or regional requirements, specific situations concerning availabilities of dates and stadiums, etc. For sake of conciseness we omit the formulation of these constraints.

A straightforward integer programming formulation of the problem could not be solved by a commercial solver after an entire day of computations. Therefore, we developed the following approach to tackle the problem. First, we add an extra constraint stating that the total number of consecutive home games and consecutive away games played by any team is fixed at its minimum. Next, the weighted sum of the broadcast objective with the penalties and incentives associated with the soft constraints is maximized, as described in the next section.

4 Solution approach

Nemhauser and Trick[6] proposed a three-phase decomposition scheme to schedule a basketball league. Our approach follows a similar strategy. In each of the three phases of the algorithm, a different subset of the problem constraints is considered.

ICMSEM Inpress

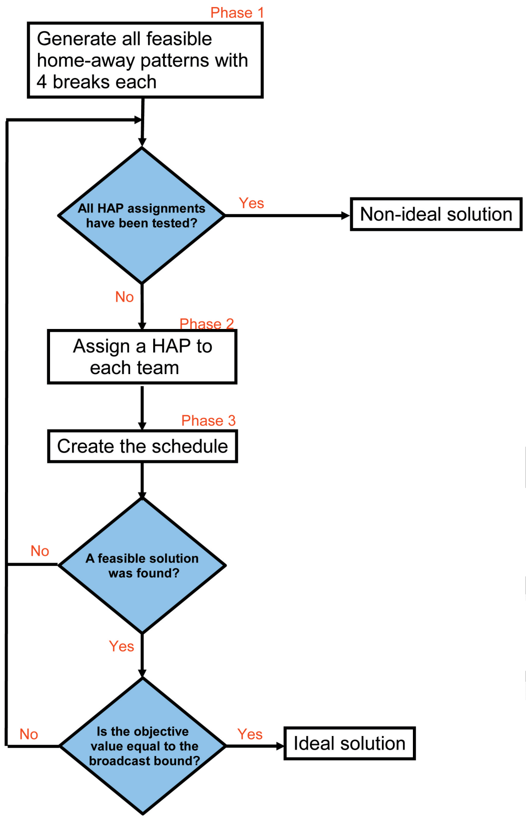

Fig. 1 Three-phase solution optimization approach

Fig. 1 depicts the three-phase solution optimization approach. The first phase is devoted to the generation of home-away patterns (HAPs) of game playing satisfying the hard constraints. Only home-away patterns with exactly four breaks are constructed, since we search for a solution minimizing the total number of breaks in which every team has the same number of breaks.

In the second phase, an explicit exhaustive enumeration scheme is used to assign a different home-away pattern of game playing to each participating team. Once a full assignment of teams to home-away patterns of game playing is obtained, the algorithm proceeds to the last phase.

Finally, in the last phase, we build and solve an integer programming problem considering the selected assignment of home-away patterns and optimizing the weighted sum objective function. This integer program defines the games to be played in each round, according with the selected home-away patterns.

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. Due to the home-away patterns assigned to each team, we also know that the number of breaks is minimum. Therefore, the number of variables is relatively small, because there is no need to use further variables to model the breaks.

In consequence, the decomposition strategy leads to a reduced model that can be solved by a commercial solver in reasonable computation times. Typically, this model may be infeasible due to the home-away pattern assignments and to the hard constraints considered at the last phase of the algorithm. If this is the case, the algorithm returns to the second phase to enumerate other, additional home-away assignment patterns and the new resulting model is solved by integer programming, until an optimal solution is found. Otherwise, the schedule is feasible and the best known solution is updated.

5 Numerical results

CPLEX 9.0 was used as the linear and integer programming solver in the computational experiments. The software is coded in C++ and runs on a standard Pentium IV processor with 256 M bytes of RAM memory.

The proposed approach was applied to the 2005 and 2006 editions of the Brazilian tournament. Ideal schedules (i.e., simultaneously optimizing both the broadcast and the fairness objectives) have been routinely obtained in less than ten minutes of running time.

First, we notice that the official schedules used in 2005 and 2006 violated some of the problem requirements, while the optimized schedules met all constraints. The feasible solutions produced by integer programming (with all soft constraints satisfied) were much better than those (non-feasible) produced by the current scheduler. Furthermore, the integer programming approach lead to schedules in which all 56 more attractive games could be broadcast by open TV channels, while the ad hoc rules used for scheduling the 2005 and 2006 tournament editions made it possible to broadcast only 43 and 47 games, respectively. These results are summarized in Tab. 1 and Tab. 2.

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

	Official schedule	Optimized schedule
All constraints	no	yes
Breaks	156	80 (optimal)
Broadcast	43	56 (optimal)

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

	Official schedule	Optimized schedule
All constraints	no	yes
Breaks	172	72 (optimal)
Broadcast	47	56 (optimal)

6 Conclusions and experience in practice

Due to the good results obtained on historical data of previous tournament editions, a software system implementing the above approach was developed and made operational through a partnership with the Brazilian Soccer Confederation.

This system was used for the first time in 2009 as the official scheduler to build the fixtures of the first and second divisions of the Brazilian national soccer tournament. Several alternative fixtures were provided to the users, who selected their preferred choice. The final fixtures obtained with this software system for both the first and second divisions of the 2009 edition of the Brazilian national soccer tournament were announced last January.

The annotated bibliography in [5] references over 150 papers connected to the use of Operations Research methods in sports scheduling. There has been a steady increase in recent years not only on the number of methodological and application papers, but also on the use of quantitative methods in practice.

This increase is powered by a large number of real-life innovative applications. In fact, sporting events generate a lot of public interest, with fairness and technical criteria to be respected and logistic issues to be optimized.

The hardness of the optimization problems in sports scheduling has lead to the use of different techniques in their solution: decomposition strategies, integer programming, column generation, Benders cuts, constraint programming, metaheuristics, and their hybrids.

The literature reports real-life applications of optimization methods to the scheduling of baseball, basketball, and hockey tournaments in the North America, as well as to soccer leagues in Austria, Brazil, Chile, Denmark, England, Germany, and Japan; to volleyball tournaments in Argentina and The Netherlands; and to the rugby World Cup. The successful work reported in this paper is one additional step towards the use of quantitative methods to improve the logistic of sporting events.

References

[1] T. Bartsch, A. Drexl, S. Kröger. Scheduling the professional soccer leagues of Austria and Germany. *Computers & Operations Research*, 2006, **33**: 1907–1937.

[2] F. Croce, D. Oliveri. Scheduling the Italian football league: An ILP-based approach. *Computers & Operations Research*, 2006, **33**: 1963–1974.

[3] G. Durán, M. Guajardo, et al. Scheduling the Chilean soccer league by integer programming. *Interfaces*, 2007, **37**: 539–552.

[4] K. Easton, G. Nemhauser, M. Trick. Sports scheduling. **in:** *Handbook of Scheduling: Algorithms, Models and Performance Analysis*, CRC Press, 2004.

[5] G. Kendall, S. Knust, et al. Scheduling in sports: An annotated bibliography. *Computers & Operations Research*. To appear.

[6] G. Nemhauser, M. Trick. Scheduling a major college basketball conference. *Operations Research*, 1998, **46**: 1–8.

[7] T. Noronha, C. Ribeiro, et al. A branch-and-cut algorithm for scheduling the highly-constrained Chilean soccer tournament. **in:** *Practice and Theory of Automated Timetabling VI*, vol. 3867, Springer, 2007, 174–186. Lecture Notes in Computer Science.

[8] R. Rasmussen. Scheduling a triple round robin tournament for the best Danish soccer league. *European Journal of Operational Research*, 2008, **185**: 795–810.

[9] R. Rasmussen, M. Trick. Round robin scheduling - A survey. *European Journal of Operational Research*, 2008, **188**: 617–636.

ICMSEM Inpress