Operations Research Techniques for Scheduling Chile's Second Division Soccer League

Guillermo Durán

Departamento de Ingeniería Industrial, Universidad de Chile, Santiago 8370439 Chile

Departamento de Matemática, FCEyN, Universidad de Buenos Aires, Buenos Aires C1428EGA, Argentina

IMAS-CONICET, Argentina, gduran@dii.uchile.cl

Mario Guajardo

Department of Finance and Management Science, NHH Norwegian School of Economics, Bergen 5045 Norway, mario.guajardo@nhh.no

Rodrigo Wolf-Yadlin

Departamento de Ingeniería Industrial, Universidad de Chile, Santiago 8370439 Chile, rwolf@dii.uchile.cl

In this paper, we use operations research (OR) techniques to schedule the Second Division of the Chilean professional soccer league. The solution must satisfy a series of conditions requested by league officials. Because the teams generally travel long distances by bus, geographical restrictions are particularly important. We specify the scheduling problem and solve it using an integer linear programming (ILP) model that defines when and where each match is played, subject to constraints. For the most difficult instances, we formulate a second ILP model that generates home-away patterns and assigns them to the teams; we then run the model, which determines the match schedule. Chilean league officials have successfully used the models to schedule all five Second Division tournaments between 2007 and 2010, replacing the random scheduling methodology that they used previously. Since 2007, the two formulations have been adapted to various formats with which the Second Division has experimented; these include a quadruple round-robin and a two-phase tournament with zonal and national phases. The application we present is one of a number of such projects that the authors and their colleagues developed over the past few years, and represents an expansion of the use of OR techniques for managing tasks in Chilean soccer.

Key words: sports scheduling; soccer; home-away patterns; integer programming.

History: This paper has been refereed.

For at least 40 years, experts in operations research (OR) and related fields have studied the application of optimization techniques to scheduling sports league seasons. Their interest has focused primarily on the difficulties of solving the large instances typical of real-world sports leagues. However, for soccer, one of the world's most popular sports, the literature on using match scheduling techniques by existing leagues has a relatively short history. In this paper, we report on the recent application of optimization to the Second Division of Chile's professional soccer league, which like the First Division is organized by the Asociación Nacional de Fútbol Profesional (ANFP or the Association).

A group that includes this paper's authors initiated the use of sports scheduling techniques in Chilean soccer for the league's First Division season of 2005. Two years later, encouraged by our good results, the ANFP requested that we extend our involvement to scheduling the Second Division season. Thus far, we have scheduled all five of the Second Division tournaments organized since our first effort at this league level in 2007.

The format of Chile's Second Division season generally differs from that of the First Division and has undergone several changes over the past four years. Therefore, we could not simply transfer the models and solution methodologies we used for the First Division to the Second. We had to adapt our techniques to the scheduling challenge and to look more deeply into the question of home-away patterns. For the more difficult instances of the Second Division schedule, we developed two linear programming models, which generate solutions in successive stages. The first model generates the clubs' (i.e., teams') home-away patterns; the second model then determines the match date assignments. This procedure has allowed us to rapidly generate solutions for all instances we encountered.

The application of our two models to the Second Division season thus marks a new approach to using OR techniques for managing tasks in Chilean soccer. It provides significant qualitative and quantitative benefits to the ANFP and the teams. These techniques have now fully replaced the traditional methodology that the ANFP used previously for Second Division scheduling.

The principal contributions of our work are threefold: first, an expansion of the use of OR in managing Chilean soccer; second, the design of a two-stage solution approach that rapidly solves the most difficult scheduling instances of the Chilean league's Second Division and that can be adapted to more efficiently generate the First Division schedule; and third, the formulation of models with the necessary versatility to provide solutions for addressing the numerous modifications made to the Second Division season format from year to year.

We structured the remainder of this paper as follows. *Previous Real-World Applications to Soccer Leagues* reviews the literature on soccer league scheduling. *Organization of Professional Soccer in Chile* explains how soccer is organized in Chile. *Scheduling the Second Division League* outlines the generic conditions and schedule formats of the Chilean Second Division over the past few years. *Solution Methods* describes the methods adopted for the Second Division scheduling problem, and *Impact and Conclusions* presents the impacts of these solutions and our conclusions. The appendices, which are in an electronic companion, show the formulation of the scheduling models. The electronic companion to this paper is available as part of the online version that can be found at http://interfaces.pubs.informs.org/ecompanion.html.

Previous Real-World Applications to Soccer Leagues

The first reported application to soccer scheduling was the work of Schreuder (1992), who scheduled the Dutch soccer association's premier league for the 1989–1990 season. His model defined the match dates for 18 teams playing each other once in a single round-robin tournament, subject to a range of commercial, sports-related, and organizational considerations. These included hard constraints that had to be strictly satisfied and soft constraints whose satisfaction was desirable but not necessary; the soft constraints were assigned weights representing their relative importance. The objective function attempted to maximize the total weight of the definitive solution. The author's solution method implemented a heuristic using construction and partitioning techniques that satisfied 80 percent of the conditions imposed for the league's 1989–1990 schedule.

Another early example was the mathematical scheduling application devised by Dubuc (Paenza 2006), which was used only once for the 1995 season of the Argentinean soccer league. The application implemented a heuristic based on simulated annealing.

Bartsch et al. (2006) reported a sports scheduling system for the German and Austrian professional soccer leagues. The 18 German league teams played a mirrored double round-robin schedule (the second half of the tournament duplicates the first, but reverses the home and away teams for each match); in Austria, 10 teams competed in a nonmirrored quadruple round-robin. In both cases, various conditions relating to attractiveness to the public, fairness, and organizational criteria were imposed; some were hard constraints and others were soft. The objective function minimized the cost of penalizing the violated constraints. The modeling approach used partially renewable resources and a heuristic solving methodology to generate the schedules. The German league used this solution approach for the 1997–1998 season and the Austrians applied it on six occasions between 1997 and 2003.

In Durán et al. (2007), the authors report the scheduling of the First Division of the Chilean professional soccer league for the 2005, 2006, and 2007 seasons (the authors have since defined the league's schedules on a permanent basis). At the time, 20 First Division clubs were classed into four groups of five and the season was divided into two phases. The first phase was a single round-robin tournament in which each team played each other team. The two top teams in the tables (standings) from each of the four groups went on to the second phase, consisting of a series of play-offs to determine the season champion. The schedule had to meet various geographical, attractiveness (to fans), and television broadcast requirements. The solution was based on integer linear programming (ILP) coupled with basic procedures for assigning each team's home-away patterns. The problem of maximizing the number of meetings between teams of the same group toward the end of the first phase was solved to optimality using branch-and-cut techniques (Noronha et al. 2007).

Rasmussen (2008) scheduled Denmark's professional league premier division for the 2006–2007 season. The 12 Danish clubs met each other three times, thus playing a triple round-robin tournament. The schedule requirements included geographical and top-team considerations expressed as both hard and soft constraints. The objective function minimized the penalties imposed for violating certain restrictions. The solution used Benders' decomposition and column-generating techniques, providing good solutions with short computation times.

Goossens and Spieksma (2009) scheduled the top-level Belgian soccer league for the 2007–2008 season. The 18 league clubs played a mirrored double round-robin tournament. The number of tournament breaks (consecutive home or away matches for a team) had to be minimized. A variety of other conditions were imposed, including television broadcast and carry-over effect requirements (i.e., team i confers a unit of carry-over effect on team j if i played the same rival club in round k that j plays in round $k+1$). The problem constraints were classified by priority level with penalties assigned. The solution approach used a mixed-integer programming (MIP) model that assigned clubs to a canonical schedule that satisfied the mirroring constraint and minimized the number of breaks. Local search procedures were then used to improve the solution by minimizing the total penalty for constraint violations.

Three additional applications have recently been reported. Flatberg et al. (2009) scheduled the Norwegian league using an integer programming model whose objective function incorporates carry-over effects. The 14 teams in this league play a nonmirrored double round-robin. Fiallos et al. (2010) scheduled the soccer league of Honduras, also using integer programming. They considered a number of conditions and assigned them different priorities. The objective function minimized the cost of penalizing the violated constraints. The 10 teams in this league play a mirrored double round-robin followed by a series of play-offs. Ribeiro and Urrutia (2011) scheduled Brazil's top league for the 2009 and 2010 seasons, maximizing the number of attractive games between elite clubs that could be broadcast on free-to-air television. They implemented an integer programming method and a three-phase solution based on a "first-break, then-schedule" decomposition scheme similar to that proposed in Nemhauser and Trick (1998). The Brazilian league has 20 teams that play a mirrored double round-robin tournament.

To the best of our knowledge, the cases outlined above are the only real-world applications of scheduling techniques to soccer leagues that have been reported in the literature. Other published studies reflect the true organization of existing leagues or some of the leagues' features, but do not report their use in practice. Of course, one or more applications might simply not be documented in academic publications. We confirmed that optimization techniques have been used for scheduling leagues in the Czech Republic since the 2002–2003 season (D. Fronceck, M. Meszka, pers. comm.), and in Poland for 2007–2008 and 2008–2009 (M. Meszka, pers. comm.).

Extensive literature on scheduling the match calendars of other sports also exists. In recent surveys, Rasmussen and Trick (2008) and Kendall et al. (2009) have thoroughly reviewed the theoretical and practical aspects of this body of work. However, considering that the Fédération Internationale de Football Association (FIFA), the world governing body for soccer, has 208 member associations, the number of scheduling applications relative to the sport's many national leagues seems small; manual scheduling techniques still predominate around the world.

Organization of Professional Soccer in Chile

Chile has only two professional soccer leagues: the First Division, which had the country's 18 strongest teams in 2010, and the Second Division, which has 14 other clubs. At the end of each season, the two top teams in the Second Division are promoted to the First Division and the last two in the First Division are relegated to the Second Division. The Second Division teams in 3rd and 4th place play against the First Division teams in 16th and 15th place, respectively, in a promotion relegation play-off; the winners play in the First Division the following year and the losers play in the Second Division.

All the teams aspire for a place in the First Division, which is the most popular league among soccer fans and the most attractive for private team sponsors and television networks. Nevertheless, the Second Division is also important because of the interest generated by the promotion possibilities from one season to the next. Furthermore, Second Division teams are usually located outside of Santiago, Chile's capital and by far its most populous city; thus, they represent smaller communities where soccer plays a central role in local entertainment. In 2010, none of the 14 Second Division teams was located in Santiago, while all 7 of Santiago's professional teams played in the First Division. The Second Division clubs are popular among local residents who tend to identify strongly with them, making them attractive to regional private sector sponsors. This has increasingly captured the interest of the television networks, which have broadcast a number of Second Division matches in recent seasons.

A principal task of the ANFP is to define the schedules for the First and Second Division soccer seasons. Until 2005, it used a system based on a predetermined canonical schedule. Each match date was paired with two numbers to which the teams were randomly assigned, automatically generating a schedule that determined the order of the matches. Although these schedules met the season's basic requirements, many teams complained because the schedules did not consider some relevant criteria.

For the 2005 season, the ANFP approached a group of OR researchers (including this paper's authors) at the University of Chile to request support in using sports scheduling techniques to develop the First Division schedule. We formulated an ILP model that incorporates the requirements of the ANFP, the teams, and the television broadcaster. These criteria include factors relating to geographic location, economics, scheduling fairness, and attractiveness to fans.

Based on the schedule's success at the First Division level, the ANFP asked us to assist in scheduling the 2007 Second Division season. Although the approach we used for the First Division was useful, the particular characteristics of the Second Division season and the special requirements of the teams meant that we needed different models to capture the specifics of the new problem. Moreover, the format of the Second Division season has undergone various changes in the past few years. The solution we developed, although still based on ILP models, adapts our First Division modeling and its underlying methodology to the Second Division's season format, number of clubs, and specific ANFP requirements. The following sections describe the principal features of this application and its evolution over the seasons.

Scheduling the Second Division League

The scheduling process for the Second Division is carried out about two months before the season begins each February. The ANFP defines in advance the dates on which the various rounds making up the schedule are to be played. Each round consists of a set of matches in which each team plays exactly once over two or three consecutive days. Weekend rounds are usually played on Fridays, Saturdays, and Sundays; midweek rounds are usually played on Tuesdays, Wednesdays, and Thursdays. Based on this calendar, we then generate a proposed assignment of matches to rounds, considering a series of conditions either requested by the teams through the ANFP, determined directly by the ANFP, or dictated by the characteristics of the season format.

Generic Conditions

In this subsection, we describe the *generic conditions* that we imposed for almost all the Second Division tournaments between 2007 and 2010, with only minimal variation. We group them using the terminology for constraint categories suggested by Rasmussen and Trick (2008).

- **Basic constraints**

  1. Each team plays a match in each round.
  2. Each pair of teams plays each other m times, the actual number depending on the type of tournament.

- **Break constraints**

  A team is said to have a break in round k if the matches it plays in rounds $(k-1)$ and k are either both at home or both away.

  3. Each team can have no more than a certain number (l_{home}) of home breaks and a certain number (l_{away}) of away breaks.
  4. In certain rounds, no team can have a break. This condition is usually imposed on the second and last rounds of a tournament for reasons of fairness (for a team

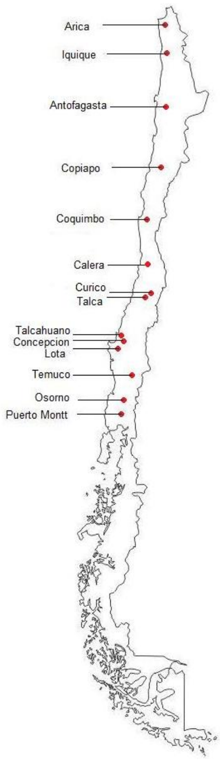

Figure 1: **The map shows the geographical location of all 14 Second Division teams in the Chilean professional soccer league in 2010.**

to start or end a tournament with two consecutive home or away games is undesirable).

- **Geographic constraints**

  Geographical aspects are usually considered in sports scheduling (Kendall 2008, Bonomo et al. 2012, Wright 1994). Because Chile stretches more than 4,300 kilometers from its northern to southern extremes (see Figure 1), a club such as Arica in the far north must travel 2,250 kilometers to play Curicó Unido in the central zone and 3,080 kilometers to visit Puerto Montt in the south.

  However, despite the distances, the teams usually journey to and from away matches by bus because air travel in Chile is expensive. Travel costs are a major expense for Second Division teams, which have limited resources for financing road trips; these trips typically include about 25 persons including players, coaching staff, medical personnel, and club officials. Thus, although teams return home immediately after an away game when the subsequent round is within a week, if a game is only three days later (e.g., a Saturday game following a Wednesday game) and the two matches are both on the road (i.e., an away break), the team might decide to travel directly to the second game for economic reasons. Generally speaking, such teams are those located at either extreme of the country. For them, and for the league as a whole, it is therefore important that the scheduling incorporate geographical restrictions, as the following constraints illustrate.

5. The number of trips teams make to relatively distant matches must be distributed evenly over the length of the tournament. Formulating this constraint involves defining clusters of teams based on their geographical locations. The precise definition will vary depending on which clubs are in the Second Division in any given season; however, we generally define three clusters representing three geographical zones of the country: North, Center, and South. Lower and upper bounds ($t_{low}^{a,b}$ and $t_{up}^{a,b}$, respectively) are set specifying the minimum and maximum number of times a team in cluster a can play a team in cluster b on a certain number s of consecutive tournament rounds.

6. If the tournament contains midweek rounds, a certain number of good trips (i.e., a trip in which a team plays away at distant locations on two consecutive rounds, one of which is midweek) must be scheduled for each team. Scheduling two such matches in a single trip provides the club with transportation cost savings that outweigh its lodging expenditures. As an example of a good trip, a team at home in the North (South) cluster might play a match on a Wednesday in the Center and another in the South (North) the preceding or following weekend.

7. Long trips on which teams play two consecutive away matches on weekends should be avoided. This restriction is formulated by imposing the restriction that if a team plays two away games in a row, at least one of them must be against a club in the same cluster. Exceptions to this rule occur with midweek games, as we

explained in the previous constraint.

- **Complementary constraints**

8. That certain pairs of teams have their home-away patterns sequenced so that one is at home while the other plays away and vice versa might be a requirement. Such team pairs are defined as complementary (or crossed). This may be desirable or necessary to coordinate the matches of two teams who share a single stadium, to maintain public order (i.e., to avoid the possibility of clashes between the fans of two teams based in the same city playing at home on the same date), or to ensure that a game is scheduled every weekend in a given region of the country.

- **Place constraints**

9. Some home-match rounds are predefined at a team's specific request. For example, a club might want to schedule a home game in the round closest to its anniversary to attract a higher turnout and thus generate additional revenue.

10. Some away match rounds are also predefined at a team's specific request. For example, this might occur if the club's stadium is booked on that date for some other event or is closed for maintenance or repairs.

- **Game constraints**

11. Some match dates are predefined at the ANFP's request. For example, it might prefer to schedule classic rivalries, which attract larger crowds, toward the end of the tournament when the stakes are higher.

- **Pattern constraints**

12. If the number of games each team plays during a tournament is even, half of the games must be played at home and the other half away. If the number is odd, the numbers of home and away games must differ by only one.

13. No team can play more than two away matches in any three consecutive rounds.

14. No team can play more than two home matches in any three consecutive rounds.

Types of Second Division Tournaments

In 2007, a change took place in the ANFP leadership, which the presidents of the Association's professional member clubs elect every four years. As a result, the formats of the Second Division season have been modified several times as the ANFP searches for a setup that is both attractive to the public and economically viable for the clubs. The widely dispersed locations of the current set of Second Division teams and their limited budgets have heavily influenced the formats adopted. Over the past four seasons, two format types have been used with variations: a twice-mirrored quadruple round-robin (used in 2007 and 2008) and a two-phase tournament in which the first phase is played in two separate geographical groups (used in 2009 and 2010). These formats differ from both the First Division season (which presently consists of a mirrored double round-robin tournament scheduled over the year) and any other system reported in the literature on soccer scheduling, as we reviewed above in *Previous Real-World Applications to Soccer Leagues.*

The format that most closely resembles the one that the Second Division used in 2007 and 2008 is Austria's quadruple round-robin; however, the latter is not mirrored. In other sports, the Finnish hockey league's season is also a nonmirrored quadruple round-robin (Kyngäs and Nurmi 2009). Formats similar to the Second Division's 2009 and 2010 tournaments can be found in the National Basketball Association, the National Hockey League, and Major League Baseball in the United States, where the teams are divided into conferences that in turn subdivide into divisions. In all these cases, the division groupings consider geographical location.

The most important features of the various Second Division tournament formats since 2007 are briefly outlined below.

The 2007 season. The Second Division had 11 clubs in 2007. The season was divided into four quarters in which each team played each other team exactly once. Because the number of teams was odd, one team had a bye in every round. The second quarter was the mirror image of the first; therefore, if team i played at home against team j on the nth round of the first quarter, team j hosted team i on the same round of the second quarter. Analogously, if team i had a bye on the nth round of the first quarter, it also had a bye on the same round of the second quarter. A symmetry condition established that the third quarter was the same as the first and the fourth the same as the second. Thus, we call the format a *twice-mirrored quadruple round-robin with an odd number of teams.* The 2007 season consisted of 220 matches distributed over 44 rounds.

The 2008 season. This season was similar to the 2007 season; however, because 12 teams (instead of 11) played, it included no byes. Its 264 matches were distributed over 44 rounds.

The 2009 season. In 2009, 14 teams played; unlike the other years, the season was broken up into two separate tournaments, the opening and the closing, each of which was a mirror image of the other. The winner of each tournament met in a play-off match. The winner moved up to the First Division, while the loser entered the promotion-relegation play-off against the 16th-place team in the First Division. The team that won the most points over the two tournaments was also promoted and the club with the second highest total entered the promotion-relegation play-off against the 15th-place team in the First Division.

The opening and closing tournaments were both divided into two phases, the zonal phase and the national phase. In the zonal phase, the 14 teams in the Second Division were divided into two groups (North Group and South Group) of 7 teams each on the basis of geographical proximity. The two groups played a single round-robin tournament in which each team played half of its matches at home and half away. Because the number of clubs in each group was odd, one team had a bye on each round. However, in the national phase, all 14 teams played a single round-robin tournament in which each played each other team. Point totals won in the zonal phase carried over into the national phase.

The ANFP asked us to add the condition that the home-away status of a match between any pair in the zonal phase would be reversed for the national phase (i.e., if team i played at home against j in the zonal phase, j would play at home against i in the national phase).

This constraint made solving the 2009 season's problem considerably more difficult than that of the other seasons, as we explained in *Solution Methods* above. In 2009, 266 matches were played over 40 rounds.

The 2010 season. In 2010, the Second Division returned to a single-tournament season incorporating certain features of the two formats used in previous years. The tournament consisted of two phases denoted as the zonal phase and the final phase, respectively. In the zonal phase, the 14 teams were split into two groups of 7, both based on geographical proximity. Each group played a twice-mirrored quadruple round-robin tournament with an odd number of teams. The top four teams in each group at the end of the tournament went on to the final phase in which the eight teams play a mirrored double round-robin, with team point totals starting over from 0. Although the identities of these teams could not be known beforehand, knowing that they will in any case be the first four teams from each group means the schedule can be generated in advance, including the geographic constraints. Thus, in addition to defining the match rounds, the final phase schedule specifies the participating teams by position, not by name, in the final zonal phase standings. For example, a game between teams $1G_1$ and $3G_2$ in the final phase means that the team that finishes at the top of group G_1 plays against the team placed in the third position in the final standings of group G_2.

The ANFP requested that this scheduling be completed prior to the season's start (i.e., together with the zonal phase because the final phase begins a few days after the zonal phase ends).

The Association also requested that the rounds in which certain teams play at home or away be coordinated with First Division matches played in the same cities to preserve a local balance of home games. The objective is similar to the complementary constraints described above, although we note that the First Division is scheduled prior to and independently of the Second. Finally, the 2010 season consisted of 224 matches played in 42 rounds.

For all the seasons 2007 through 2010, the goal was to define schedules that satisfy all the requested constraints. Thus, the central problem is one of feasibility rather than optimization. Even so, the large number of constraints (numbering in the thousands) that must be satisfied made manually generating good schedules practically impossible. Indeed, the pre-2007 scheduling did not fulfill many of the constraints that have since been imposed (e.g., those relating to geographic conditions).

In the 2010 season, the ANFP asked that, in addition to satisfying all of the conditions, we concentrate on scheduling as many attractive matches (i.e., matches that pair off the top two teams in each zonal group) as possible in the last three rounds of the final phase. We formulated an objective function to incorporate this goal into our modeling (see the formulation of the objective function in Appendix A in the electronic companion to this paper).

Solution Methods

We specified the Second Division scheduling problem using an ILP model that we call the match scheduling model (MSM). Given two teams and a round, the main decision variable is a binary one, which equals one if these teams play each other in that round. In tournaments with an odd number of teams, byes are expressed as a match against a dummy team added to the model during the implementation (the scheduling conditions for this team include only the basic constraints). Appendix A includes a detailed description of the model, which we implemented and solved using the CPLEX solver package.

Scheduling the twice-mirrored quadruple round-robin tournaments for 2007 and 2008 was relatively straightforward. In a single run, the MSM found a feasible solution in a few minutes. This was possible for two reasons. First, the number of teams (11 or 12) was limited and the size of the problem therefore relatively small; second, the mirror and symmetry conditions described above significantly reduced the search space. Once a schedule for the first quarter is generated, the schedules for the entire tournament can be determined. For example, a condition for the fifth round in the second quarter can be expressed in terms of the fifth round in the first quarter, its mirror image, because of the mirror condition. Similarly, a condition on a given round in the third quarter can be given in terms of its analogous round in the first quarter because of the symmetry condition. The relationship is analogous for rounds in the second and fourth quarters.

The 2010 season was also fairly easy to schedule. The characteristics of the tournament were such that the scheduling of the zonal phase could be separated into two independent subproblems both of which were relatively small because each group had only seven teams. The final phase was scheduled separately from the zonal phase and was less restrictive given that the teams involved were not known a priori. The size of the problem for the final phase was also relatively small given that only eight teams participated. Solution to optimality using the objective function defined in the previous section (i.e., maximizing the number of attractive matches in the last three rounds) was achieved in a matter of seconds.

The two 2009 tournaments were more difficult to define because of the linkage between the zonal and national phases. We considered various alternatives for approaching the problem. The first alternative was the most obvious—to solve the zonal and national phases simultaneously. In this alternative, the scale of the integer model was large—about 5,000 binary variables and 2,500 constraints.

Unfortunately, when we used this approach, we could not find feasible solutions through direct application of the MSM, either alone or together with the basic pattern-assignment procedures used for the First Division, as the computational section in Durán et al. (2007) discusses.

Another alternative was to divide the problem into two smaller subproblems, one for each phase. We could then further decompose the zonal phase subproblem into two components, one for each group of teams (North and South). The MSM for each group had approximately 500 binary variables and 500 constraints, whereas the national phase subproblem had about 3,000 binary variables and 1,500 constraints.

The disadvantage of this decomposition approach is that the subproblems are not independent of each other, because they are linked by three groups of constraints: first, the reversal of the home-away status of matches between any pair of teams i and j that meet in both phases; second, the maximum number of allowed breaks per team (note that a home (away) match on the last game of the zonal phase and again on the first game of the national phase counts as a home (away) break); and third, the condition that a team cannot play three consecutive home or away matches (including cases similar to the one just described in which a team ends the zonal phase with a home (away) break and starts the national phase with another home (away) match). The decomposition into two subproblems also necessarily implies that one of the two phases must be solved first; the possibility exists that the solution generated will lead to the infeasibility of the second subproblem, although the whole problem is not necessarily infeasible.

We were also unable to obtain a feasible solution to the problem using this decomposition with the MSM (or variations of it, applying basic pattern-assignment procedures). Therefore, we had to develop another ILP model for generating the home-away patterns. Appendix B (see the electronic companion to this paper) shows this formulation, which we call the pattern generating model (PGM).

Using the two models (PGM and MSM), we attempted to schedule the season in two stages. In the first stage, the PGM constructed and assigned each team's home-away patterns applying some of the problem constraints. In the second stage, we used the MSM to determine which teams played each other in each round, with all of the constraints satisfied. Because the PGM already fulfilled most of the conditions, such as the complementary and pattern constraints, we could simplify the MSM runs by omitting these restrictions. This technique of breaking down the problem into stages is common in sports scheduling, although the precise implementation methods vary; Rasmussen and Trick (2008, Sect. 4) review this strategy.

The addition of the PGM to the solution also provided an opportunity to incorporate another key consideration: the creation of a set of patterns at an early point in the solution. Nemhauser and Trick (1998) identified this as critical to successful schedule generation. We believe that it gives the PGM a distinct advantage over other formulations found in the literature. For each team, the PGM constructs individual patterns that consider team-specific constraints from the categories described in the *Generic Conditions* subsection.

If the MSM does not achieve feasibility with the pattern assignment from the PGM, then the MSM is executed again using the same set of patterns; however, only heavily constrained teams (typically 50 percent of the teams) are preassigned patterns, leaving the MSM to determine the assignments for the remaining teams. In most cases, these two iterations will result in the MSM obtaining a feasible schedule with the initial set of patterns from the PGM; if not, we run the PGM again to generate a new set of patterns, which the MSM then applies. In practice, we found that if the constraints of the PGM are reordered, CPLEX will generally deliver a different solution. If not, some constraints can be applied to eliminate some patterns from the solution so that the PGM is forced to generate a different set. The whole procedure may be repeated as necessary until a feasible schedule is attained.

Using this two-model approach, solving the zonal and national phases simultaneously produced a feasible solution in about 10 minutes; solving them separately took about 30 minutes. This execution time for the simultaneous solution was satisfactory for practical purposes; therefore, we adopted this mechanism to generate the various schedules that the ANFP requested for the 2009 season. We note here that the actual scheduling is an interactive process undertaken in conjunction with Association officials and extends over at least two weeks. During this period, various new conditions, which the officials or the teams had originally considered, are incorporated. Thus, the models must be run several times before the definitive version of the season schedule is determined.

In light of the advantages derived from implementing the PGM for the Second Division, we developed a variation on this specification to generate patterns for the First Division schedules. We then implemented a variation on the MSM to determine the definitive version. Using this two-stage process reduced the total solution time to 10 percent of the time previously needed to generate the First Division schedule. One reason was that we constructed the patterns by considering, from the start, the teams to which they would be assigned. Another factor was that the number of patterns generated was the same as the number of teams; previously, we usually had to generate more patterns than teams.

Impact and Conclusions

The schedules generated by the methods presented above for the five Second Division tournaments played between 2007 and 2010 have had positive qualitative and quantitative impacts. Although we cannot directly measure the quantitative impact or isolate it from exogenous effects, we can nevertheless makes some observations.

The financial savings to teams based in the far north and south of the country, which stem from scheduling good trips, are particularly relevant. Each tournament scheduled using our methods included approximately five such trips (except in 2009 when all rounds were played on weekends and the good-trip notion thus did not apply). This compares favourably with the last manually generated schedules (2005 and 2006), which had no good trips. The ANFP has estimated that each of these trips represents a savings to the associated team of almost $4,000—about 13 percent of a Second Division team's monthly payroll, which averages $30,000, and 5 percent of the $75,000 in annual gate receipts of clubs such as Puerto Montt in the far south.

Another important impact is the more even distribution of trips over the length of the tournament, enabling the teams to better spread their expenses over the year, thus balancing the monthly resources of the clubs, especially those with tight budgets who are particularly dependent on home-game revenues and sponsorship contracts paid on a monthly basis. By contrast, in the last manually generated schedules, teams from the far ends of the country were assigned various double sequences of long Sunday-to-Sunday away trips.

The application of the OR models also greatly facilitated implementation of the key changes that the ANFP made to the 2009 and 2010 seasons. These modifications were designed to ensure attractive tournaments for fans, while reducing road trips from one end of the country to the other, easing the strain on the team budgets and alleviating player travel fatigue.

We also note that the Second Division's 2007 season, the first to involve us in scheduling this league, saw an increase of about 10 percent in average match attendance over the previous season when the traditional method of random scheduling was still used. Although OR techniques cannot take credit for all of this increase, we believe it was a major contributor.

The classic match between Rangers de Talca and Curicó Unido, which was set for the last round of each season quarter in 2007, is an example. The ANFP requested this match; it felt that the match would be more attractive if played late in the quarter when it would likely be more decisive. At the end of the season, the Rangers had a good chance at promotion to the First Division; thus, the possibility of clinching it at home against their classic rival made the encounter especially interesting to fans. The game drew 10,006 spectators, well above the average attendance of 2,761 for the other Rangers matches that season. Gate receipts were \$57,600, more than seven times the \$8,000 the club averaged for home games that year. It was the second highest gate in the entire league in 2007. It is highly unlikely that the random date assignment method would have made a similar contribution to fan interest and team finances.

The qualitative impacts of the OR scheduling since 2007 have taken various forms. Because the requirements incorporated into the models are known to all the teams, the process of generating the schedules is more transparent. In addition, the ANFP accepts requests for the incorporation of requirements from all the teams, thus making the system fairer. The ability of our models to be adapted to the various changes in the Second Division season over the past few years demonstrates the flexibility of our approach.

Appropriate scheduling of matches through the application of OR methodologies has also made the tournaments more attractive, to the approval of the fans. The effect is particularly noticeable in classic matches, for which both club-supporter enthusiasm and player motivation have noticeably increased. Future research by the authors on the various Second Division tournament formats will include the analysis of fan preferences and team results, both financial and on the field. Goossens et al. (2010) offers an insightful analysis of soccer season formats in the context of the Belgian league; in the Chilean case, the frequent changes, as we describe in the *Types of Second Division Tournaments* subsection, offer an interesting opportunity to study different formats, especially because ANFP management has still not settled on an ideal season format.

Our modeling techniques have also contributed to maintaining public order by including criterion such as not scheduling home games for teams for dates on which their home town is hosting other large events. For example, round 4 of the 2009 tournament coincided with a major music festival, which required a large police presence, in the city of Viña del Mar. Therefore, we imposed a place constraint that none of the four teams that play in the Viña del Mar area would have a home game that day.

The ANFP, the teams, the fans, and the press have generally welcomed our application of OR scheduling, ensuring its credibility in Chilean soccer circles. Since we extended our scheduling to the Second Division in 2007, the Association has received no complaints from the teams regarding match schedules of this league.

A key aspect of our scheduling efforts for the Chilean soccer league has been our alliance with ANFP management. In the weeks before each tournament, we work closely with management to define the conditions to be incorporated as model constraints. From the beginning, ANFP has understood that OR models provide valuable support for the decision making involved in schedule design and that our methodology has enabled it to include considerations that the traditional scheduling procedure could not accommodate, thus increasing its control over the process. We view this understanding, which has been fundamental to the continued success of the alliance, as one of the most difficult achievements for sports scheduling initiatives, perhaps even more difficult than solving the scheduling problems. In preliminary contacts with other sports leagues, the unwillingness of decision makers to consider any changes has been the main barrier encountered in adopting new scheduling methodologies. In stark contrast to such attitudes, ANFP management has told us that it could not imagine scheduling its tournaments today without the help of OR models.

The application of our scheduling models to the Second Division came about as a natural extension of the success we had in developing models in conjunction with the ANFP for the First Division, all 14 of whose tournaments we scheduled between 2005 and 2011. In the words of Association president Harold Mayne-Nicholls, "We are highly satisfied with these supporting models for game scheduling and expect to continue using both the First and Second Division systems." As this article is being written, the authors are in the process of scheduling the Second Division's 2011 season; its format will be similar to that of the 2009 season described in the *Types of Second Division Tournaments* subsection.

The use of OR techniques has also spread to other areas of Chilean football. Sports scheduling tools were used to define the matches of three of Chile's youth league seasons between 2005 and 2010. These techniques were also used to design a proposed schedule that the ANFP submitted in 2007 to the South American Football Confederation for the South American qualifying tournament of the 2010 World Cup. The positive results with these projects have prompted us to join with the ANFP in developing an optimization model that would generate fairer and more efficient referee assignments for soccer matches. The model was tested in the 2010 Second Division tournament and in the youth league season (Alarcón et al. 2009, 2010).

Our experience with sports scheduling in Chile also led to applications in neighboring Argentina. Since 2007, one of the authors has participated in scheduling its professional volleyball league using similar tools (Bonomo et al. 2012), and the authors of that work have begun discussions with the governing bodies for Argentinean soccer and basketball.

In conclusion, this study provides further evidence of the potential for constructing solid and lasting partnerships between academia and the public and private sectors. These partnerships generate significant benefits for both sides.

Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://interfaces.pubs.informs.org/ecompanion.html.

Acknowledgments

We thank the *Asociación Nacional de Fútbol Profesional de Chile*, and in particular its president, Harold Mayne-Nicholls, and its competition manager, René Rosas. This project would not have been possible without their support. We also thank Andres Weintraub who encouraged us to write this article, Fernando Alarcón and Kenneth Rivkin for their many useful suggestions, and the three anonymous reviewers and the Special Issue Associate Editor whose comments noticeably improved the final version. Finally, we are grateful to the Complex Engineering Systems unit of the Millennium Sciences Institute (Chile) for financial support in the completion of this work. The first author was partially supported by ANPCyT Grant PICT-2007-00518 (Argentina), UBACyT Grant 20020100100980 (Argentina), and FONDECyT Grant 1110797 (Chile).

References

Alarcón, F., G. Durán, M. Guajardo. 2010. Assigning referees to a Chilean football tournament by integer programming and patterns. B. McCollum, E. Burke, G. White, eds. *Proc. 8th Internat. Conf. Practice Theory Automated Timetabling (PATAT 2010)*, 466–467.

Alarcón, F., G. Durán, M. Guajardo. 2009. Un modelo de asignación de árbitros para el torneo de fútbol chileno y un enfoque de resolución en base a patrones (in Spanish). *Revista Ingeniería de Sistemas* **23**(1) 125–143.

Bartsch, T., A. Drexl, S. Kröger. 2006. Scheduling the professional soccer leagues of Austria and Germany. *Comput. Oper. Res.* **33**(7) 1907–1937.

Bonomo, F., A. Burzyn, A. Cardemil, G. Durán, J. Marenco. 2012. An application of the traveling tournament problem: The Argentine volleyball league. *Interfaces*

Durán, G., M. Guajardo, J. Miranda, D. Sauré, S. Souyris, A. Weintraub, R. Wolf. 2007. Scheduling the Chilean soccer league by integer programming. *Interfaces* **37**(6) 539–552.

Fiallos, J., J. Pérez, F. Sabillón, M. Licona. 2010. Scheduling soccer league of Honduras using integer programming. A. Johnson, J. Miller, eds. *Proc. 2010 Indust. Engrg. Res. Conf.* Cancún, MX.

Flatberg, T., E. J. Nilssen, M. Stølevik. 2009. Scheduling the topmost football leagues of Norway. Accessed September 29, 2011, http://folk.uio.no/trulsf/pub/euro2009.pdf.

Goossens, D., F. Spieksma. 2009. Scheduling the Belgian soccer league. *Interfaces* **39**(2) 109–118.

Goossens, D., J. Beliën, F. Spieksma. 2010. Comparing league formats with respect to match importance in Belgian football. *Ann. Oper. Res.* (June 15) 1–18.

Kendall, G. 2008. Scheduling English football fixtures over holiday periods. *J. Oper. Res. Soc.* **59**(6) 743–755.

Kendall, G., S. Knust, C. C. Ribeiro, S. Urrutia. 2009. Scheduling in sports: An annotated bibliography. *Comput. Oper. Res.* **37**(1) 1–19.

Kyngäs, J., K. Nurmi. 2009. Scheduling the Finnish major ice hockey league. *Computational Intelligence in Scheduling, CI-Sched '09. IEEE Symposium* 84–89.

Nemhauser, G. L., M. A. Trick. 1998. Scheduling a major college basketball conference. *Oper. Res.* **46**(1) 1–8.

Noronha, T. F., C. C. Ribeiro, G. Durán, S. Souyris, A. Weintraub. 2007. A branch-and-cut algorithm for scheduling the highly-constrained Chilean soccer tournament. *Lecture Notes in Computer Science* Vol. 3867, 174–186.

Paenza, A. 2006. Matemática...Estás Ahí? Episodio 2 (in Spanish). *Siglo XXI*, Buenos Aires, Argentina, 194–206.

Rasmussen, R. V. 2008. Scheduling a triple round robin tournament for the best Danish soccer league. *Eur. J. Oper. Res.* **185**(2) 795–810.

Rasmussen, R. V., M. A. Trick. 2008. Round robin scheduling - A survey. *Eur. J. Oper. Res.* **188**(3) 617–636.

Ribeiro, C., S. Urrutia. 2011. Scheduling the Brazilian soccer tournament: Solution approach and practice. *Interfaces*, Forthcoming.

Schreuder, J. 1992. Combinatorial aspects of construction of competition Dutch professional football leagues. *Discrete Appl. Math.* **35**(3) 301–312.

Wright, M. 1994. Timetabling county cricket fixtures using a form of tabu search. *J. Oper. Res. Soc.* **45**(7) 758–770.

Author Queries

A1 Au: Production: the volume and page numbers must be added.