Scheduling the Chilean Soccer League by Integer Programming

Article *in* Interfaces · December 2007

DOI: 10.1287/inte.1070.0318 · Source: DBLP

INTERFACES
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0092-2102 | EISSN 1526-551X | 00 | 0000 | 0001

INFORMS
DOI 10.1287/xxxx.0000.0000

Scheduling the Chilean Soccer League by Integer Programming

Guillermo Durán, Mario Guajardo, Jaime Miranda,
Denis Sauré, Sebastián Souyris, Andrés Weintraub
Department of Industrial Engineering, Faculty of Physical and Mathematical Sciences, University of Chile,
(gduran,maguajar,jmiranda,dsaure,ssouyris,aweintra)@dii.uchile.cl, www.dii.uchile.cl

Since 2005, Chile's professional soccer league has used a game-scheduling system based on an integer linear programming model. The Chilean league managers have considered several criteria for the last tournaments' scheduling, involving operational, economic and sporting factors, thus generating a highly constrained problem, in practice unsolvable by their last methodology. This led to the adoption of a model with real conditions, some of them totally new in the use of sports scheduling techniques in soccer leagues. The schedules so obtained have meant greater benefits for the teams, given by lower costs and higher incomes, fairer seasons and tournaments that are more attractive to sports fans. Such success has completely fulfilled the expectations of the *Asociación Nacional de Fútbol Profesional (ANFP)*, the organizing body for Chilean professional soccer.

Key words: Chilean soccer league, integer programming, sports scheduling

1. Introduction

Soccer is *"the passion of multitudes"* around the world, a phenomenon that was amply demonstrated by the last World Cup held in Germany. But beyond the purely sporting and emotional aspects of the game, its management increasingly requires the application of scientific criteria. In Chile, soccer has been subject to even more competition not only from the international leagues and other televised sports but also from new types of activities and better access to the existing ones such as shopping malls, cinema, video games, the Internet, and so on. The organizers of other professional sports are up against similar situations in various countries.

This scenario has brought about a falling-off in Chileans' interest in soccer, which has translated into a drop in revenue generated by the sport. Professional league officials find themselves facing the challenge of boosting the attractiveness of the league season in the hope of reversing this decline and also reducing costs. One of the most important instruments for achieving this is the planning of league game schedules. The task of scheduling each regular-season matchup taking into account the many different factors that would ensure a game calendar that is simultaneously fair to the teams, economically beneficial and attractive to sports fans would be nearly impossible if attempted manually.

Beginning with the 2005 season, the *Asociación Nacional de Fútbol Profesional (ANFP)*, the organizing body for Chilean soccer, has employed the services of the *Centro de Gestión de Operaciones (CGO)*, a unit of the Industrial Engineering Department at the *University of Chile*, to assist in the planning of the league's game schedule. We integrate sporting, operational and economic criteria within an integer programming model to come up with a schedule that meets the criteria established by the *ANFP* and makes the season more interesting for soccer fans.

The nature of the programming involved falls within the area known as sports scheduling. In this paper we present the criteria used in defining the efficiency of a season schedule in terms of sporting fairness or equity, the introduction of operational and economic considerations into the scheduling process, and how the model currently employed and its implementation lend the process a degree of flexibility that was previously absent in Chilean soccer schedules. In addition to increasing the season's attractiveness, these factors combine to put the scheduling process on a more scientific basis, making it more transparent and therefore more acceptable to team managers.

The paper is organized as follows. We start with a description of the Chilean soccer tournaments and a review of the literature on sports scheduling. Then, in section 3, we explain the conditions considered on the problem. In section 4, we refer to the model and its computational solution. In section 5, we show some recent figures and qualitative factors which have satisfied the *ANFP*, teams, TV and fans, with the application of our model. Conclusions and guidelines for future work are discussed in section 6. Finally, the formulation of the mathematical model is specified in the Appendix.

2. Background

The First Division of Chile's soccer league contains 20 teams and divides its annual playing calendar into two halves, known respectively as the Opening Championship and the Closing Championship. Each of the championships, also called tournaments, is in turn comprised of two phases: the regular season, consisting of 19 playing dates known as rounds, and the playoffs. The teams are organized into 4 groups of 5 teams each, and each team must play once against each of the other 19 (a simple round-robin system). Both the date of each round and the composition of the groups are set beforehand by the *ANFP*. Once the regular season is over, the two top teams in each group advance to the playoffs where a champion is decided. This setup was inspired by the Mexican soccer league system.

As opposed to typical US sport tournaments, the Chilean soccer league is composed by 3 divisions. Each year the last two teams (measured as the sum of points in both tournaments) of the First Division are relegated to the Second Division, while the two top teams of the Second Division are promoted to the First Division (and similarly occurs with the Second Division and the Third Division).

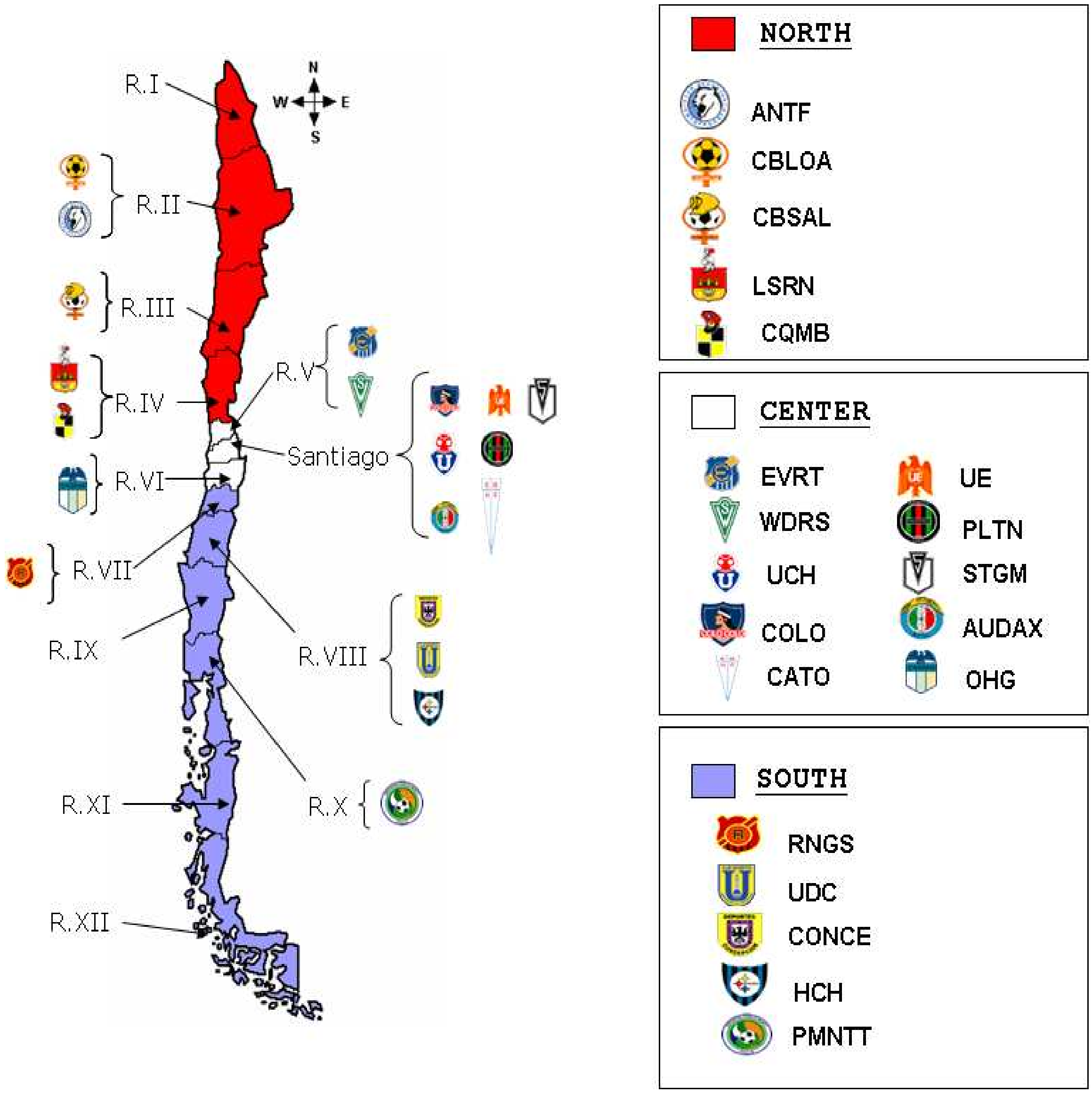
It is important to remark that the country is geographically divided in 12 Regions and the Metropolitan Region (Santiago), which is located between Regions V and VI. We have classified the 20 teams of the First Division in three clusters by geographic location: North with 5 teams, Center with 10 teams and South with 5 teams (see Figure 1).

Previous to 2005 the schedule of the First Division had been decided by a random draw of teams and venues in a preset template, as it is done in almost all soccer leagues in South America and Europe. With this system the season could readily be scheduled manually, but no account was taken of the majority of criteria a schedule could reasonably be expected to fulfill in order to be *"efficient"*.

To begin with, the schedule should be *"fair"* in sporting terms. This means, among other things, that each team should play a balanced mix of home and away games against the strongest teams, games against the strongest teams should not be scheduled consecutively, and each team plays against two of its group opponents at home and against the other two opponents away.

Second, certain economic and operational considerations that would mean greater revenue and lower costs for the teams should be incorporated. For example, scheduling two consecutive away games (Sunday-Wednesday, or Wednesday-Sunday) for a given team in different opponents' venues located relatively close to each other but far from the team's home venue would constitute a *"good"* trip in that it would spare the team a second long trip. Other examples would be setting *"attractive"* games for appropriate dates, such as summer home games for teams located in popular beach towns; and scheduling classic rivalries or matches between teams of the same group in the second half of the tournament when the stakes are higher. Also, we can distribute weekday home games fairly given that such dates are less attractive to the teams because attendance is lower than on weekends (revenue for any specific game go entirely to the home team).

Figure 1 Map of Chile with the location of the 20 teams of the First Division (in the Opening Tournament 2006).



Scheduling in previous years on these criteria was woefully deficient: classic matchups on inappropriate dates, weaker teams playing all their games against stronger ones away from home, unbalanced distribution of weekday home games, etc.

A season calendar that met the above-cited standards of efficiency would be nearly impossible to devise manually, and it is precisely here that operations research can make a contribution, applying management technology that would flexibilize and automate the season scheduling process. To appreciate the scale of the complexities involved, we need merely note that for a tournament in which 6 teams play a simple round-robin there are 720 different possible schedules (even without considering whether games are at home or away), while for a tournament with 8 teams there are more than 30 million possibilities. Clearly, the number of possible schedules for the 20 soccer teams in Chile's First Division would be simply unimaginable.

The use of sports scheduling techniques is still a novelty in South America, the only known previous case being the Argentinean soccer league which used such a method developed by E. Dubuc for the 1995 season (see (12)), before abandoning the practice. Isolated cases have also been reported in European soccer (3). In the United States, on the other hand, sports scheduling is routinely employed by the most important basketball, baseball and football leagues, which maintain their own teams of academics or hire third-party companies to design efficient regular-season schedules (see, for example, (4), (11) and (10); also, Pittsburgh Business Times of November 12th., 2004 (13) and Associated Press, December 1st., 2004 (2)).

At the academic level, the literature on sports scheduling has grown significantly in recent years. Various articles have been published proposing as-yet untried scheduling applications for existing leagues ((6), (7), (15)).

Interest in the problem increased notably with the publication of the *Traveling Tournament Problem (TTP)* (8), which involves designing a schedule that minimizes the distances teams in a sports league must travel. Though not set up as a real case, the TTP has generated a significant benchmark using a range of methods and algorithms. In (1) a heuristic is proposed for solving the TTP based on simulated annealing, while (5) presents a tabu search application for the same purpose. In (9) a combination of integer programming and constraint programming is offered as a method of finding the optimal solution for leagues of up to 8 teams, and in (14) heuristics are developed for the mirrored version of the TTP (double round-robin tournament in which the game order in each single round-robin is the same).

As we will describe later, the goal of the Chilean tournament is not precisely minimizing traveled distances, but rather finding a schedule satisfying a long and complex list of conditions which combine to produce to a highly constrained problem.

3. Conditions imposed on the problem

In what follows we describe the conditions that must be met by the schedule to comply with the requirements established by the *ANFP* for the 2006 Opening Tournament. Most of them were also among the criteria applied for the two 2005 tournaments.

- **Basic schedule constraints**

1. Each team plays each of the others once over the course of the 19 rounds in the tournament.
2. Each team plays each round either at home or away.
3. Each team plays at least 9 rounds, but not more than 10, at home.

- **Home and away game sequence constraints**

4. Each team plays at most one sequence of two consecutive rounds at home (home stand). This condition implies that no team plays more than 2 consecutive rounds at home.

5. Each team plays at most one sequence of two consecutive rounds away (road trip). This condition implies that no team plays more than 2 consecutive rounds away.

6. Let A be a set of rounds denoted *"adjustment rounds"*. If a team plays at home (away) in an adjustment round, it must play away (at home) in the following round. In the 2006 Opening, the set was defined as $A = \{1, 16, 18\}$. This is intended to balance teams' home and away games between the early and late stages of the tournament.

- **Home game balance constraint for matches against group rivals**

7. Each team plays against two of its group opponents at home and against the other two opponents away.

• **Geographic constraints for double away game sequences**

Certain constraints have been incorporated to avoid consecutive long road trips.

8. When a North (South) team plays 2 consecutive away games, neither of them will be in the South (North).

9. When a North (South) team plays 2 consecutive away games, at least one of the games will be in the North (South).

10. When a Center team plays 2 consecutive away games, at least one of the games will be in the Center.

Is is important to note that, unlike 2005, in the 2006 Opening Tournament all the rounds were played on weekends (and it is very expensive for a team to stay away from its city a whole week). So, we could not schedule good trips in consecutive away games Sunday-Wednesday or Wednesday-Sunday.

• **Constraints on highly popular teams (Colo Colo, Universidad Católica, Universidad de Chile)**

11. If a team plays at home (away) against Colo Colo, it plays away (at home) against Universidad de Chile. This contributes both to fairness and a better balance of revenue between the Opening and Closing tournaments given that games against these teams generally generate greater receipts.

12. The 3 classic matchups between these popular teams must be played between the $10th$ and $16th$ rounds (these parameters may change from tournament to tournament).

13. Each of the 3 popular teams plays exactly one classic matchup at home.

Note that these three teams (Colo Colo, Universidad Católica, Universidad de Chile) are called in the appendix, figure and tables as COLO, CATO and UCH, respectively.

• **Operational constraints on the availability of mobile broadcasting units for televising games**

Four games from each round are televised. These include all of the ones involving the three popular teams plus a fourth match which is the one whose two opposing teams have the highest combined point total going into the game (if there is a classic game in a given round, two games are selected in this way). Given Chile's long geographical extension it is desirable that these teams do not play in venues located far apart, thereby limiting travel distances and the associated costs faced by the television broadcasting company for the transfer of mobile units to and from the games.

14. When a popular team plays in the North (South), neither of the 2 other popular teams can play in the South (North).

15. Given that the first 5 rounds are scheduled for the middle of the Summer (when many events are televised and the availability of mobile units is lower), none of the 3 popular teams play away on those rounds in outlying areas of the country. Outlying areas are defined as the members of a set containing the teams whose home venues are located north of Chile's Region IV or south of Region VIII.

• **Constraints on strong teams (the 3 popular teams plus Cobreloa)**

16. No team may play 2 consecutive games against a strong team.

17. Games between Cobreloa (CBLOA), a fourth strong team, and the popular teams are played between the 6*th* and the 18*th* rounds.

- **Constraints on home and away games for "crossed" teams**

A pair of teams are referred to as *"crossed"* if their home venues are in the same region and if, for operational reasons (e.g., they share the same home stadium), security reasons or in order not to leave any region without a match for an entire weekend, they alternate with each other playing home and away games in each round. In total, there were 5 crossed pairs: Wanderers (WDRS) and Everton (EVRT), both teams of the Region V, or Colo Colo and Universidad de Chile, the most important teams in the country both located in Santiago, are examples of crossed pairs.

18. When one team of a crossed pair is playing at home, the other team plays away, and vice versa.

- **Constraints on regional classic matchups**

A number of pairs of teams from the same region with a historic rivalry are defined as a set of regional classics.

19. Regional classic matchups are held between the 8*th* and 18*th* rounds.

- **Constraints on Santiago games**

20. The number of games held in Santiago in each round cannot be less than 2 or more than 4 (there are 7 Santiago teams). This enables the amount of soccer activity in the capital to be regulated and ensures the availability of stadiums and municipal security personnel.

21. It is desirable that the 4 Santiago teams with lowest drawing-power do not play against each other in the first 5 rounds (all in summer) as the attendance would be relatively low. A set D consisting of these 4 teams is defined.

- **Tourism-related constraints**

A set T is defined containing teams located in tourist areas, where it is desirable that at least one attractive game (against a popular team) be scheduled for the first rounds of the tournament, during the summer.

22. Each team located in a tourist area plays at home against at least one of the popular teams in one of the first 5 rounds.

- **Special constraints**

23. Some teams do not have their playing fields ready in time for the beginning of the tournament, and therefore should be scheduled for away games in the first round. These teams form a set S.

24. Not more than 3 games between teams of the same group are held in the last round. This condition was required by the *ANFP* in 2006, given that in the past tournaments the playoff qualifiers for most of the groups were defined before the very last round.

25. Each North (South) team shall play at least once at home against a North (South) team. This constraint will ensure that, in the 2006 Closing Tournament, North and South teams have at least one opponent in their respective clusters to play an away game against. It could help to

avoid *"bad"* two-game road trips. (Note that for Center teams, the same condition is guaranteed by constraints 11 and 13, because the three highly popular teams are from the Center cluster).

Further constraints related to special circumstances have also been incorporated, such as not scheduling home games for a team on dates when its stadium is booked for other events and avoiding road trips for teams close to dates when they have to travel to an international cup game. Also, we scheduled in the first round of a tournament a match between teams which are playing in the playoff finals of the immediately preceding championship. So, this game could be rescheduled at a later date to allow these two teams an additional week's rest (otherwise the *ANFP* would have to reschedule two games).

Since the 2005 tournaments included weekday games, other factors were taken into account relating to economic or fairness considerations. In the 2005 Opening, which only included one round played on Wednesday, it was requested that a *"good"* road trip be scheduled for at least 3 teams between that round and an immediately adjacent one. As for the Closing, which had 3 Wednesday rounds in the schedule, each team was required to play at home on at most 2 of them (clearly, in these two tournaments, constraints 8, 9 and 10 were not considered for consecutive rounds including weekday games).

Finally, for each of the championships an objective function was formulated that maximized the concentration of *"decisive"* games among the final rounds of the schedule. In the Opening Tournaments of 2005 and 2006, decisive games were defined as those between teams in the same group. In the 2005 Closing, this definition was broadened to include the games between teams that were expected, based on their performance in the Opening Tournament, to be fighting against relegation to the Second Division. The detailed formulation of the objective function is given in the Appendix.

Note should be taken here of the important role played by the iterative process involving the group of academic experts and the *ANFP* in arriving at a definitive schedule. Three or four different versions of the schedule may bounce back and forth between them as various details are fine-tuned and new special constraints are incorporated into the model before a set of final proposals are put forward. This process is carried out starting one month before the publication of the definitive schedule.

4. The mathematical model and the computational solution

The conditions described in the previous section were expressed in terms of an integer linear programming model (see Appendix). Once the model was built, *CPLEX* 9.0 was used to solve it on a Pentium IV computer with a 2.4 GHz processor.

The problem we face is essentially one of feasibility. Our main goal is to find a schedule that meets all of the conditions imposed on it. The objective function simply measures how much we can *"push"* the decisive games towards the end of the tournament. Clearly, then, it is not so important that we arrive at an optimal solution, unlike most other optimization problems. The *ANFP* selects the final schedule from among a series of options presented to it by us (after the iterative process we mentioned above), and the choice it makes is not necessarily the one that performs best according to the objective function.

The principal family of decision variables in the model, set out in detail in the Appendix, takes on a value of 1 when team i plays at home against team j in round k, and 0, otherwise. Other variables express some of the more complex conditions. The model as formulated is an extremely hard to solve problem. To simplify the solution, an additional factor was incorporated that establishes certain home-away patterns for the teams, adding some rigidity to the formulation.

A home-away pattern (see, for example, (9) and (11)) is a sequence assigned to a given team that indicates the number of rounds it will play at home and away. A sequence consisting of the

19 elements $(H, A, ..., H)$ represents a pattern in which H signifies a home game round and A an away game round. The constraints applied to the assignment of home-away patterns are given in the Appendix.

The advantage of this approach with home-away patterns is the reduction in solution time achieved without significant loss of solution quality. Furthermore, it can ensure a priori that certain constraints which normally account for the most serious prolongations of model run time are satisfied. Two such constraints are those regarding crossed pairs of teams and the avoidance of double home stands or road trips.

For the 2005 Opening Tournament a feasible solution was found using constraint programming (*CP*) after approximately three hours of run time. The model solved was similar to the one given here in the Appendix but without the objective function, being simply a feasibility problem. The task was performed by *ILOG's* Solver 5.2 software. *CP* was employed because of its good performance solving similar sports scheduling problems in other cases (e.g., (10)).

Once this solution was completed, the set of patterns so obtained was assigned fixedly to each team. The optimization process was executed and the optimum for this pattern assignment was obtained in a matter of minutes. The schedule approved for this tournament featured 24 of the 40 games between teams of the same group in the last three rounds, much to the satisfaction of the *ANFP*. Note that it is the maximum possible number for 3 rounds, because at most 8 games between teams of the same group can be played in a given round.

For the 2005 Closing Tournament the procedure adopted was similar, except that the assignment of *"feasible"* patterns was partially incorporated into the optimization process. In order to keep solution time down to a reasonable level, some patterns were assigned by the process while others were held fixed. It was discovered through experimentation that with up to about 15 fixed patterns and 5 to be assigned by the process, the solution was still arrived at quickly, requiring not more than 15 minutes, and was better than the solution obtained using the method employed for the Opening. Thus, the calendar approved for the 2005 Closing scheduled for rounds in the second half of the tournament the great majority of games between teams from the same group and between teams which were expected to be fighting one other to avoid relegation to the Second Division. A Closing Tournament is significantly more constrained, because the home condition for a given pair of teams is already defined as the opposite of the one scheduled for the Opening Tournament (i.e., if team i played at home against team j in the Opening, then j must play at home against i in the Closing).

For the 2006 Opening Tournament, a feasible solution was also sought as a first step. This involved imposing a set of patterns derived from the contents of the definitive solution for the 2005 Closing, with a certain number of modifications introduced that would eliminate all double home stands and road trips and impose several different sequences for rounds in which a concentration of intergroup games was desired. Once these patterns were imposed, a feasible solution was found in just a couple of minutes. The set of patterns in this solution was then taken but instead of being assigned to the teams as is, they were completely incorporated into the optimization process, based on the solution of the relaxed problem. We solved the LP problem (that is, without the constraint of integrality of the variables) and then set to one the *"pattern"* variable z_{ip} with highest value in the LP solution. We repeated this procedure sequentially up to get a feasible solution. When the model did not generate a feasible result, we backtracked through the iterations and changed the pattern assignment in a logical way. This heuristic guided us to good solutions, better than the original we had, in a short time, as the LP problem can be solved in a matter of seconds. Once the 20 patterns were fixed we proceeded as with the 2005 Closing, applying this solution to the model as the initial solution while allowing a few patterns (between two and five) to be reassigned in an attempt to further improve the objective function value.

The solution chosen by the *ANFP* for the 2006 Opening Tournament is shown in Figure 2. In this schedule, 100% of the games between teams in a given group were set for the $10th$ or later

Figure 2 **Schedule approved for the 2006 Opening Tournament. Games between teams of the same group are shown in grey. Note that all of them are scheduled for the second half of the Championship.**

Teams \ Match	1	2	3	4	5	6	7	8	9
UCH	UE	@RNGS	UDC	@WDRS	AUDAX	@CBSAL	PMNTT	@ANTF	PLTN
COLO	@OHG	HCH	@LSRN	EVRT	@PLTN	WDRS	@CONCE	CBSAL	@STGM
CATO	@EVRT	PMNTT	UE	@CQMB	ANTF	@CBLOA	@PLTN	LSRN	@OHG
ANTF	@HCH	PLTN	@EVRT	STGM	@CATO	OHG	@LSRN	UCH	@CONCE
CBLOA	@CBSAL	CQMB	@STGM	CONCE	@OHG	CATO	@HCH	RNGS	@WDRS
CBSAL	CBLOA	@LSRN	@OHG	PMNTT	@STGM	UCH	EVRT	@COLO	CQMB
LSRN	@CONCE	CBSAL	COLO	@RNGS	UE	@AUDAX	ANTF	@CATO	HCH
CQMB	UDC	@CBLOA	@PLTN	CATO	@PMNTT	HCH	@WDRS	OHG	@CBSAL
EVRT	CATO	@AUDAX	ANTF	@COLO	@HCH	CONCE	@CBSAL	PLTN	@PMNTT
WDRS	@PLTN	OHG	@CONCE	UCH	RNGS	@COLO	CQMB	@UE	CBLOA
AUDAX	@PMNTT	EVRT	@HCH	OHG	@UCH	LSRN	@STGM	CONCE	@RNGS
UE	@UCH	CONCE	@CATO	HCH	@LSRN	PLTN	@RNGS	WDRS	UDC
PLTN	WDRS	@ANTF	CQMB	@UDC	COLO	@UE	CATO	@EVRT	@UCH
STGM	RNGS	@UDC	CBLOA	@ANTF	CBSAL	@PMNTT	AUDAX	@HCH	COLO
OHG	COLO	@WDRS	CBSAL	@AUDAX	CBLOA	@ANTF	UDC	@CQMB	CATO
RNGS	@STGM	UCH	@PMNTT	LSRN	@WDRS	@UDC	UE	@CBLOA	AUDAX
UDC	@CQMB	STGM	@UCH	PLTN	@CONCE	RNGS	@OHG	PMNTT	@UE
CONCE	LSRN	@UE	WDRS	@CBLOA	UDC	@EVRT	COLO	@AUDAX	ANTF
HCH	ANTF	@COLO	AUDAX	@UE	EVRT	@CQMB	CBLOA	STGM	@LSRN
PMNTT	AUDAX	@CATO	RNGS	@CBSAL	CQMB	STGM	@UCH	@UDC	EVRT

Teams \ Match	10	11	12	13	14	15	16	17	18	19
UCH	@HCH	COLO	OHG	@CQMB	STGM	@EVRT	@CATO	LSRN	@CBLOA	CONCE
COLO	RNGS	@UCH	@AUDAX	CATO	@UE	CBLOA	CQMB	@PMNTT	ANTF	@UDC
CATO	CONCE	@STGM	RNGS	@COLO	WDRS	@AUDAX	UCH	@UDC	CBSAL	@HCH
ANTF	WDRS	@RNGS	CBSAL	@UDC	AUDAX	@PMNTT	CBLOA	@CQMB	@COLO	UE
CBLOA	LSRN	@UDC	UE	@PLTN	PMNTT	@COLO	@ANTF	AUDAX	UCH	@EVRT
CBSAL	@UDC	PLTN	@ANTF	HCH	@RNGS	WDRS	@UE	CONCE	@CATO	AUDAX
LSRN	@CBLOA	@OHG	WDRS	@PMNTT	UDC	@STGM	EVRT	@UCH	CQMB	@PLTN
CQMB	STGM	AUDAX	@CONCE	UCH	@EVRT	UE	@COLO	ANTF	@LSRN	RNGS
EVRT	OHG	@UE	UDC	@STGM	CQMB	UCH	@LSRN	WDRS	@RNGS	CBLOA
WDRS	@ANTF	HCH	@LSRN	AUDAX	@CATO	@CBSAL	UDC	@EVRT	STGM	@PMNTT
AUDAX	UE	@CQMB	COLO	@WDRS	@ANTF	CATO	PLTN	@CBLOA	UDC	@CBSAL
UE	@AUDAX	EVRT	@CBLOA	@OHG	COLO	@CQMB	CBSAL	@STGM	PMNTT	@ANTF
PLTN	PMNTT	@CBSAL	STGM	CBLOA	@HCH	CONCE	@AUDAX	RNGS	@OHG	LSRN
STGM	@CQMB	CATO	@PLTN	EVRT	@UCH	LSRN	@CONCE	UE	@WDRS	OHG
OHG	@EVRT	LSRN	@UCH	UE	@CONCE	RNGS	PMNTT	@HCH	PLTN	@STGM
RNGS	@COLO	ANTF	@CATO	CONCE	CBSAL	@OHG	HCH	@PLTN	EVRT	@CQMB
UDC	CBSAL	CBLOA	@EVRT	ANTF	@LSRN	HCH	@WDRS	CATO	@AUDAX	COLO
CONCE	@CATO	@PMNTT	CQMB	@RNGS	OHG	@PLTN	STGM	@CBSAL	HCH	@UCH
HCH	UCH	@WDRS	PMNTT	@CBSAL	PLTN	@UDC	@RNGS	OHG	@CONCE	CATO
PMNTT	@PLTN	CONCE	@HCH	LSRN	@CBLOA	ANTF	@OHG	COLO	@UE	WDRS

rounds, and a bit more than 80% were concentrated between the $14th$ and the $19th$ rounds, the ones most highly weighted in the objective function. The value of the feasible solution in the objective function was 607, with a gap of 5.6% separating it from the relaxed problem solution (taking into account the last cut shown in the Appendix).

5. Results

It is not easy to measure the impact of using the described system because many factors other than scheduling influence on variables such as attendance at stadiums.

Nevertheless, certain observations can be made with confidence. In both 2004 tournaments, the last under the old system, the classic matchup between Universidad de Chile and Colo Colo was held in the 1st round and drew crowds of 26,000 people in the Opening and 22,000 in the Closing. By contrast, in the two 2005 tournaments and the 2006 Opening this matchup was scheduled further into the season and the attendance figures jumped to 45,000, 37,000 and 49,000, respectively.

Two further indicators of interest are the attendance and revenue averages, both measured per game, during the regular season portions of each Championship (the playoffs remaining unaffected by the new scheduling criteria). Table 1 summarizes the data for the last two tournaments under the old scheduling system (2004 Opening and Closing) and the latest tournament played (2006 Opening). They reveal that the rise in attendance average was 32% compared to the 2004 Opening and 39% over that year's Closing, whereas the corresponding increases for ticket sales average were 102% and 94%, respectively.

Table 2 compares attendance and ticket revenue for the classic games of the 2004 tournaments (still using the old scheduling method) with those of the 2005 and 2006 tournaments scheduled under the new system. When the same team was at home, all of the 2005 matchups had better crowds and revenue than 2004, with total attendance up 74% and receipts up 142%. As for the 2006 Opening classics (once again, with the same team at home), they drew 124% more than their 2004 equivalents translating into a revenue rise of 347%.

Table 1 **Comparisons of average game attendance and ticket revenue (in thousands of Chilean pesos) in the 2004 Opening and Closing Tournaments and the 2006 Opening Tournament. Figures refer only to regular season games. (Data supplied by ANFP.)**

Averages Attendance and Ticket Revenue per Game			
	2004 Opening	2004 Closing	2006 Opening
Attendance	3,756	3,557	4,953
Ticket Revenue	5,852	6,095	11,803

Table 2 **Comparisons of attendance and ticket revenue (in thousands of Chilean pesos) for classic games in 2004, 2005 and 2006. Figures refer to matches in which the home team was the same. (Data supplied by ANFP.)**

Attendance and Ticket Revenue for Classic Games					
	2004 Op.	2005 Op.	2004 Cl.	2005 Cl.	2006 Op.
	UCH @ COLO		COLO @ UCH		
Date	Feb-08-04	Apr-10-05	Aug-01-04	Aug-28-05	Apr-09-06
Attendance	25,743	45,236	21,750	37,420	48,996
Ticket Revenue	55,900	114,879	59,967	137,394	240,557
	2004 Cl.	2005 Op.	2004 Op.	2005 Cl.	2006 Op.
	CATO @ UCH		UCH @ CATO		
Date	Nov-04-04	Apr-30-05	May-04-04	Nov-13-05	May-14-06
Attendance	18,093	24,450	7,881	18,292	14,409
Ticket Revenue	55,173	71,499	12,241	69,099	67,466
	2004 Op.	2005 Op.	2004 Cl.	2005 Cl.	2006 Op.
	COLO @ CATO		CATO @ COLO		
Date	Mar-06-04	Mar-20-05	Sep-12-04	Sep-25-05	Apr-23-06
Attendance	9,887	24,352	13,333	18,138	32,654
Ticket Revenue	16,575	100,408	29,595	61,074	147,442

Regarding tourist areas, Table 3 displays comparisons for home games played in summer rounds by teams in Region IV, Coquimbo (CQMB) and La Serena (LSRN), and Region V, Everton (EVRT) and Wanderers (WDRS), against the popular teams. According to the data, attendance grew by 46% for Region IV matches between 2004, when they were not played in summer, and 2006, when they were. The improvement in Region V was particularly impressive at 156%. In both cases revenue also increased significantly, by 84% and 313% respectively.

Furthermore, the adoption of criteria that reduces the costs of broadcasting games has been highly celebrated by the TV managers and the own *ANFP*, which has the 80% share of the soccer TV company. For example, two televised games played in the north may result in savings for the television company of around $US\$20,000$ over the cost of broadcasting one game held in the north and the other in the south. At least part of these savings could be passed in future to the teams. In fact, the income from TV rights is a very important financial source for the Chilean teams. In the last negotiation the league sold its TV rights for approximately $US\$3,700,000$. Forty percent of this amount was assigned to the three most popular teams (U. Chile, Colo-Colo, U. Católica).

It should be kept in mind that though the quantitative information in the preceding paragraphs may be explained in part by the new schedules, exogenous factors difficult to control for are also present that may distort the measurements. These include the performance of the national team

Table 3 **Comparisons of attendance and ticket revenue (in thousands of Chilean pesos) for games held in tourist areas in 2004, 2005 and 2006. Figures refer to matches in which the home team was the same in both years. (Data supplied by ANFP.)**

Attendance and Ticket Revenue for Games in Tourist Areas				
Region IV	2004	2006 Op.	2004	2006 Op.
	COLO @ LSRN		CATO @ CQMB	
Date	May-09-04	Feb-11-06	Mar-13-04	Feb-18-06
Attendance	5,373	7,533	4,673	7,178
Ticket Revenue	10,175	22,101	10,681	16,175
Region V	2004	2006 Op.	2004	2006 Op.
	CATO @ EVRT		UCH @ WDRS	
Date	Apr-21-04	Jan-29-06	Oct-29-04	Feb-19-06
Attendance	3,314	6,638	3,494	10,787
Ticket Revenue	3,840	13,857	6,563	29,091

in World Cup qualifying rounds, weather patterns, acts of violence committed by team supporters and the quality of the teams (fundamentally those that draw larger crowds).

With regard to qualitative factors, the positive impact of the sporting fairness criteria adopted in our model deserves special mention. Both the *ANFP* and the league teams have expressed their satisfaction with the way these criteria have been put into practice.

Another qualitative aspect noted by senior *ANFP* officials is that for the first time in many years they have received almost no complaints from the teams regarding the schedules.

6. Conclusions

In order to improve the solution process, the procedures described above for finding good solutions raise certain issues to be explored in future work. These include:

- Formalization of the solution to the integer problem starting from the relaxed version via a heuristic that sequentially structures the fixing of the z_{ip} pattern variables at 1, detecting and repairing infeasibilities.
- Incorporation of the creation of patterns in the optimization process or use of a set of patterns of higher cardinality than the number of teams. This could lead to better solutions and greater flexibility in the search for feasible solutions, but likely at the price of greater solution times.
- Experimentation to determine how much the cuts shown in the Appendix contribute in terms of reducing solution time (recall that as with every integer programming problem, the use of cuts to adjust the feasible polyhedron of the linear relaxation may be highly useful).

The incorporation of these modern techniques into Chilean soccer league scheduling process since 2005 has provided an excellent opportunity to demonstrate that the use of Operations Research can be effective in making soccer season schedules more attractive to the public as well as fairer and more profitable for the teams and organizing bodies.

It is also important to underline the transparency brought to the system by the model under discussion here. Once the constraints to be applied are defined and made known to all concerned, they are incorporated as part of the mathematical model. Then we can generate some possible solutions among which the *ANFP* will choose the definitive version. This procedure also requires the *ANFP* to submit its objectives for the schedule to the teams, which facilitates consensus-building and the creation of new mechanisms for improving the league's scheduling process.

There are several anecdotes that clearly illustrate the impact created by the scheduling application in Chilean soccer circles, one of which is particularly revealing. The day the 2005 Closing Tournament schedule was made public, the *ANFP's* operational manager noted that he had received a complaint from the president of the Coquimbo team regarding its home game against Colo Colo. It had been scheduled on a local holiday known as the "Fiesta de la Pampilla", which meant that the necessary police presence for the match would not be available as well as the attendance at stadium would be significatively reduced. Within a few minutes the game was switched to a different round, and since the overall schedule had already been published the number of other changes were minimized so that the modified version was as similar as possible to the original one. Crucial to the task was the ability to rely on a rapid and agile tool that would enable us to solve the problem in very little time.

Other significant reactions have been expressed to us by some of the actors involved, such as the comment made by a player for Palestino, one of Santiago's team, who was surprised by the *"chance occurrence"* that his team was scheduled to play against its four group rivals in the last five rounds of the 2005 Opening, which led to more significant games. Another was the television company executive who remarked during the 2006 Opening on their *"luck"* at having to televise games involving Universidad de Chile on a Saturday and Universidad Católica the next day while the two teams were on road trips in the north, thus saving the company a significant amount due to the proximity of the two venues and the correspondingly low mobile unit transfer costs.

In conclusion, it is worth observing that management techniques can make other contributions to South America's most popular sport. Issues such as resource management, managing the creation of lower divisions, new tournament formats, optimal ticket prices, strategic alliances with other countries in the region, policies for encouraging the return of top players currently playing abroad, and the efficient administration of the economic and operational aspects of teams and related organizations, are just some of the areas that could benefit from a more scientific and quantitative approach.

Acknowledgments
To both anonymous referees for the valuable suggestions which improved this work. We would also like to thank the Asociación Nacional de Futbol Profesional de Chile, and in particular Alejandro Carmash, the *ANFP's* operational manager, and Felipe Chaigneau, its executive secretary, without whom this project would not have been possible. Finally, we are grateful to the Complex Engineering Systems unit of the Millennium Sciences Nucleus for financial support in the completion of this concrete. The first author was also partially financed by *FONDECYT* grant no. 1050747, Conicyt, Chile and *UBACyT X*184, Universidad de Buenos Aires, Argentina. The first and last authors received partial funding from *PROSUL* 490333/2004 − 4, CNPq, Brazil.

References

[1] A. Anagnostopoulos, L. Michel, P. Van Hentenryck, and Y. Vergados. A simulated annealing approach to the traveling tournament problem. *Journal of Scheduling*, 9:177–193, 2006.

[2] Associated Press (December 1st., 2004). Husband-wife team out bid after 24 years. (http://sports.espn.go.com/mlb/news/story?id=1936328&CMP=OTC-DT9705204233).

[3] T. Bartsch, A. Drexl, and S. Kröger. Scheduling the professional soccer leagues of Austria and Germany. *Computers and Operations Research*, 33(7):1907–1937, 2006.

[4] J. Bean and J. Birge. Reducing traveling costs and player fatigue in the National Basketball Association. *Data Mining and Knowledge Discovery*, 10:98–102, 1980.

[5] A. Cardemil and G. Durán. Un algoritmo tabú search para el traveling tournament problem (in Spanish). *Revista Ingeniería de Sistemas*, 18(1):95–115, 2004.

[6] D. Costa. An evolutionary tabu-search algorithm and the NHL scheduling problem. *INFOR*, 33:161–178, 1995.

[7] F. Della Croce and D. Oliveri. Scheduling the Italian football league: an ILP-based approach. *Computers and Operations Research*, 33(7):1963–1974, 2006.

[8] K. Easton, G. Nemhauser, and M. Trick. The traveling tournament problem: description and benchmarks. In *Proceedings of the 7th. International Conference on Principles and Practice of Constraint Programming*, pages 580–584, Paphos, 2001.

[9] K. Easton, G. Nemhauser, and M. Trick. Solving the travelling tournament problem: a combined integer programming and constraint programming approach. In E. Burke and P. De Causmaecker, editors, *PATAT 2002, Lecture Notes in Computer Science*, volume 2740, pages 100–109. Springer, 2003.

[10] M. Henz. Scheduling a major college basketball conference-revisited. *Operations Research*, 49:163–168, 2001.

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

[12] A. Paenza, *Matemática... ¿Estás ahí? Episodio 2* (in Spanish), pages 194–206. Siglo XXI, Buenos Aires, Argentina, 2006.

[13] Pittsburgh Business Times (November 12th., 2004). Striking it big: Local firm lands major league baseball scheduling contract.
(http://www.bizjournals.com/industries/sports_business/general/2004/11/15/pittsburgh_story7.html).

[14] C.C. Ribeiro and S. Urrutia. Heuristics for the mirrored traveling tournament problem. *To appear in European Journal of Operational Research*, 2006.

[15] J. Schreuder. Combinatorial aspects of construction of competition Dutch professional football leagues. *Discrete Applied Mathematics*, 35:301–312, 1992.

Appendix. Formulation of Mathematical Model

The integer linear programming model used to generate the 2006 Opening Tournament is described below.

A. Variables

To define the games to be held in each round, we define $\forall i \neq j \in I$ (the set of teams) and $\forall k \in K$ (the set of rounds), a family of binary variables, as follows:

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

To represent simply certain home and away game sequence constraints, we define $\forall i \in I$ and $\forall k = 1, \ldots, 18$ the following auxiliary variables, also binary:

$$y_{ik} = \begin{cases} 1 & \text{if team } i \text{ plays at home in rounds } k \text{ and } k+1. \\ 0 & \text{otherwise.} \end{cases}$$

$$w_{ik} = \begin{cases} 1 & \text{if team } i \text{ plays away in rounds } k \text{ and } k+1. \\ 0 & \text{otherwise.} \end{cases}$$

A total of more than 7,900 variables are included.

B. Objective Function

To conform with *ANFP* requirements, our objective function maximized the concentration of games between teams in the same group towards the final rounds of the tournament.

The experience of previous tournaments suggests it is not advisable to overload the last round, as by that time the playoff qualifying teams tend already to be determined. More specifically, it was requested that the majority of games involving teams of the same group are concentrated between the 14*th* and 19*th* rounds, that not more than 3 of these games be scheduled for the last round, and that, to the extent possible, none of them be played before the 10*th* round.

With the foregoing in mind, a game between teams of the same group played on the final round was assigned a weight of 15, while all other games held after the 9*th* round were assigned a weight equal to the round number. Thus, the objective function used in the model was the following:

$$\max \left\{ \sum_{10 \leq k \leq 18} \sum_{e} \sum_{i \in t(e)} \sum_{j \in t(e)} k \cdot x_{ijk} + \sum_{e} \sum_{i \in t(e)} \sum_{j \in t(e)} 15 \cdot x_{ij19} \right\},$$

where:

$t(e)$ denotes the set of teams in group e, $e \in E = \{1, 2, 3, 4\}$.

In the 2005 Opening Tournament, the objective function incorporated a weighted sum of the round numbers for the games between teams of the same group. In the 2005 Closing Tournament, games between relegation rivals were also weighted (a set of 6 candidates was deduced by the performance of the teams in the Opening Tournament).

C. Constraints

The formulation of the constraints is given below. They are numbered to match the numbering of the constraints described in Section 3.

1. $\sum_k [x_{ijk} + x_{jik}] = 1 \qquad \forall\, i, j \in I$
2. $\sum_j [x_{ijk} + x_{jik}] = 1 \qquad \forall\, i \in I,\, k \in K$
3. $9 \leq \sum_j \sum_k x_{ijk} \leq 10 \qquad \forall\, i \in I$
4. $\sum_{k<19} y_{ik} \leq 1 \qquad \forall\, i \in I$

To this must be added the constraint relating the y variables to the x variables:

$$\sum_j [x_{ijk} + x_{ij(k+1)}] \leq 1 + y_{ik} \qquad \forall\, i \in I,\, k < 19$$

5. $\sum_{k<19} w_{ik} \leq 1 \qquad \forall\, i \in I$

To this must be added the constraint relating the w variables to the x variables:

$$\sum_j [x_{jik} + x_{ji(k+1)}] \leq 1 + w_{ik} \qquad \forall\, i \in I,\, k < 19$$

6. $\sum_j \left[x_{ijk} + x_{ij(k+1)}\right] = 1 \qquad \forall\, i \in I,\, k \in A$
7. $\sum_k \sum_{j\, \in\, t(e)} x_{ijk} = 2 \qquad \forall e \in E,\, i \in t(e)$
8. $\sum_{i \in South} [x_{ijk} + x_{ij(k+1)}] \leq 1 - w_{jk} \qquad \forall j \in North,\, k < 19$

$$\sum_{i \in North} [x_{ijk} + x_{ij(k+1)}] \leq 1 - w_{jk} \qquad \forall j \in South,\, k < 19$$

9. $w_{ik} \leq \sum_{j \in North} [x_{jik} + x_{ji(k+1)}] \quad \forall i \in North,\, k < 19$

$$w_{ik} \leq \sum_{j \in South} [x_{jik} + x_{ji(k+1)}] \quad \forall i \in South,\, k < 19$$

10. $w_{ik} \leq \sum_{j \in Center} [x_{jik} + x_{ji(k+1)}] \quad \forall i \in Center,\, k < 19$
11. $\sum_k [x_{hik} + x_{hjk}] = 1 \quad \forall\, h \neq i, h \neq j,\; i = COLO, j = UCH$
12. $\sum_{i,j \in PopularTeams} \sum_{(10>k\, \vee k>16)} x_{ijk} = 0$
13. $\sum_k [x_{hik} + x_{jik}] = \sum_k [x_{hjk} + x_{ijk}] \qquad h = CATO, i = COLO, j = UCH$
14. $\sum_{i \in North} x_{ijk} \leq 1 - \sum_{i \in South} x_{ihk} \qquad \forall j, h \in PopularTeams, j \neq h, \quad k \in K$
15. $\sum_{k \leq 5} \sum_{j \in OutlyingTeams} x_{jik} = 0 \quad \forall i \in PopularTeams$
16. $\sum_{j \in StrongTeams} \left[x_{ijk} + x_{jik} + x_{ij(k+1)} + x_{ji(k+1)}\right] \leq 1 \qquad \forall\, i \in I,\, k < 19$
17. $\sum_{j \in PopularTeams} \sum_{(6>k\, \vee k>18)} [x_{ijk} + x_{jik}] = 0 \qquad i = CBLOA$
18. $\sum_h [x_{ihk} + x_{jhk}] = 1 \qquad \forall\, (i,j) \in CrossedTeams,\, k \in K$
19. $\sum_{(8>k\, \vee k>18)} [x_{ijk} + x_{jik}] = 0 \qquad \forall\, (i,j) \in RegionalClassics$

20. $2 \leq \sum_{i \in Santiago} \sum_{j} x_{ijk} \leq 4 \qquad \forall k \in K$

21. $\sum_{6>k} x_{ijk} + x_{jik} = 0 \qquad \forall\, i,j \in D$

22. $\sum_{j \in PopularTeams} \sum_{k \leq 5} x_{ijk} \geq 1 \qquad \forall i \in TouristTeams$

23. $\sum_{i} x_{ijk} = 1 \qquad \forall j \in S,\ k = 1$

24. $\sum_{e} \sum_{i \in t(e)} \sum_{j \in t(e)} x_{ijk} \leq 3,\ k = 19$

25. $\sum_{k} \sum_{j \in North} x_{ijk} \geq 1 \qquad \forall i \in North$

$\sum_{k} \sum_{j \in South} x_{ijk} \geq 1 \qquad \forall i \in South$

In total, the model considers around $3,000$ constraints.

Now let us examine certain additional aspects of formulating the constraints on patterns. We shall call P a set of 20 patterns and Home(k) the subset of the set of patterns that assigns home games in round k. Consider the following variables:

$$z_{ip} = \begin{cases} 1 \text{ if pattern } p \text{ is assigned to team } i. \\ 0 \qquad\qquad\qquad \text{otherwise.} \end{cases}$$

The constraints on the patterns are as follows:

26. Exactly one pattern is assigned to each team.
$\sum_{p} z_{ip} = 1 \qquad \forall i \in I$

27. Exactly one team is assigned to each pattern. This constraint is useful to strengthen the relaxation of the integer problem and holds only if the number of patterns in P is equal to 20. Alternatively, one might consider more patterns in this set. This might lead to better solutions but at the expense of higher running times.
$\sum_{i} z_{ip} = 1 \qquad \forall p \in P$

28. A team plays at home in a given round if the assigned pattern so indicates; otherwise, it plays away.
$\sum_{j} x_{ijk} = \sum_{p \in Home(k)} z_{ip} \qquad \forall i \in I, k \in K$

Given the complexity of the solution, we assume that it may be useful to incorporate cuts that help reduce the size of the feasible domain of the relaxed problem.

29. Note that since each group contains 5 teams, there can be at most 2 games between teams of the same group for each group and round. We therefore add the following cut:
$\sum_{i \in t(e)} \sum_{j \in t(e)} x_{jik} \leq 2 \qquad \forall e \in E,\ k \geq 10$

Finally, we add the following upper bound for the objective function. Note that this condition is implied by constraints 1, 24 and 29, but its inclusion showed to be efficient in our experiments.

30. Given that the total number of games between teams of the same group is 40, the value attainable by the objective function is upper bounded by 643:
$\sum_{10 \leq k \leq 18} \sum_{e} \sum_{i \in t(e)} \sum_{j \in t(e)} k \cdot x_{ijk} + \sum_{e} \sum_{i \in t(e)} \sum_{j \in t(e)} 15 \cdot x_{ij19} \leq 3 \cdot 15 + 8 \cdot 18 + 8 \cdot 17 + 8 \cdot 16 + 8 \cdot 15 + 5 \cdot 14 = 643$