Scheduling the games of the Portuguese major football league and referee assignment based on integer programming

André Ricardo Reis
IST
Student number: 65671
Email address: *andre.ricardo.reis@tecnico.ulisboa.pt*

October 24, 2015

Abstract

This article aims to explore the problematic of game scheduling and referee assignment for the case of the first Portuguese football league. The final objective is to provide a computational tool capable of getting good solutions for both situations. Nowadays, each of those two tasks is developed by an independent organization. FPF[1] takes care of the referee's assignment while *Liga de Clubes NOS*[2] is responsible for the game scheduling. Both problems have proper restrictions based on real requirements and will be modelled and solved recurring to binary integer programming. The solutions obtained by solving them were provided by two different software. *CPLEX* was used to get feasible scheduling solutions (schedules that satisfy all the restrictions used without considering any criteria) and also to assign referees to those games. Then, simultaneously, was used *Wolfram Mathematica* to treat the non linearity of the objective function of the scheduling model and implement it, in order to optimize the model and obtain better solutions. The objective function of the scheduling model will consider two different criteria where the first is related to the minimization of the number of consecutive home, or away, games of a team, *breaks*, and the second intends to minimize the consecutive games of smaller clubs against the biggest competitors. This dissertation will compare the results of both criteria used in the real timetable of the season 2014-2015 with the ones obtained by the model developed, and will also create and analyse two different approaches used to assign referees. The results obtained show that the proposed models optimise game scheduling and the referee assignment evidence that the actual models could be improved, with positive effects on competitiveness and fairness of portuguese championship. Those were presented to a FPF's responsible whose reaction was very positive, when considering this work as a case study exercise.

Index Terms – ***Sports Scheduling, Portuguese Football, Referee Assignment, Operations Research, Integer Programming, Binary Integer Programming***

1 Introduction

Football is globally considered as the world's most popular sport. Recalde et al. (2013) refer to a study made by *FIFA*[3] which states that 270 million people- approximately 4% of the world's population- actively participate in football. From an economic perspective, football is a source of employment with big flows of money circling. It gathers many interests from many different stakeholders, like football teams, players, sponsors, broadcasters and corporations, among others.

Clubs are fed by revenues, and those can urge, for instance, from merchandising, television and radio transmissions or from gate receipts, (Bartsch et al., 2006). That economic performance of industry players is directly affected by the design of the schedule. Designing a schedule capable of satisfying interests and wills of the agents involved in this rising industry and to provide exciting matches for the fans throughout the competition, contributing to the increase of the audience is extremely important.

The Portuguese championship architecture is named by the existing literature as a double round robin tournament, which means that it is a tournament with two rounds and, essentially, must meet basic and especial conditions intrinsic to the competition. Many of them may be difficult to achieve and that is why Operations Research is an important tool than can be used to help providing good solutions. A wide literature on sports' scheduling shows great progress towards achieving good results. Dur et al. (2005) have scheduled the Chilean football league by using integer linear programming and the model developed was successfully applied to the league in the year of 2006. Bartsch et al. (2006) applied mixed integer linear programming and an heuristic approach to solve the scheduling problem in both leagues of Germany and Austria. Della Croce and Oliveri (2006) used a three phase constructive heuristic approach, based on integer programming, to schedule the Italian first football league tournament (*Series A*) and Rasmussen (2008) scheduled the triple round robin Danish football league. Even championships with traditionally lower interest, from audiences and economic agents, like the Norwich football league, have been studied by Flatberg and Nilssen (2009) with the objective of analysing the called *carry-over* effect. This study anal-

[1] *Federação Portuguesa de Futebol*. For further information about this organization consult the following website: *http://www.fpf.pt/*

[2] For further information about this organization you can consult the following website: *http://www.ligaportugal.pt/*

[3] *Fédération Internationale de Football Association*. For further information about this organization consult the following website: *http://www.fifa.com/*

yses the effect that a team has on another when a third team plays consecutively with both of them. Fiallos et al. (2010) developed an informatics system that automatically provides good schedules to the Honduras football league, improving the previous, less scientific, manual process. Recalde et al. (2013) treated the problematic of the scheduling of the Ecuadorian football league using, again, a three phase heuristic approach combined with integer programming. Finally Ribeiro and Urrutia (2014) have studied the scheduling of the Brazilian football tournament with the objectives of minimizing the number of *breaks* and maximizing the broadcasts in open channels.

Related to the referee assignment's problematic there are two main studies that deserve special attention, due to their important contribution to this research area and were studied by Yavuz et al. (2008) and Alarcón et al. (2012). They both analysed ways of assigning referees to football games using integer programming. At last, Kendall et al. (2010) did a brilliant synthesis about all the important studies developed in the past 40 years, until the date of their publication. They refer definitions, formulations, types of problems and techniques for solving both problematic object of study.

This study will present two different integer models. One will be developed for finding a schedule for the Portuguese league respecting all the fundamental restrictions, and the other will be used to get two different solutions, by applying two different approaches, for the referee assignment. It is believed that both models will contribute to improve competitiveness and fairness of the league, since the current situation might be prone to introduce eventual competitive distortions due to external competition interests. The first approach will assign referees by raffle, and not by nomination, to all the games of the season, at the beginning of it. The second, more realistic, will make it on a weekly basis during the competition. Nowadays the referee assignment is made by nomination, on a weekly basis and it depends, among other factors, on the previously match performance of each referee and his weekly physical and psychological evaluation. This second approach will raffle referees essentially based on their match performance classifications.

This paper is organised as follows. In section 2 the problems are defined. In section 3 both integer models are formulated, with the presentation of some needed notation for that case and, in section 4, some computational results and respective analysis are shown. Finally, in section 5, the conclusions, comments and observations will be drawn and also some suggestions for further research steps will be provided. This last chapter will finish with the realization of the very deserved acknowledgements.

2 Problem Definition

This section defines the problem inherent to this research by presenting all the restrictions that nowadays are used by both organizations evolved in the competition. Those restrictions were directly gathered in reunion with them. Due to that, they can be considered as official information, for the 2014-15 football season.

2.1 The Portuguese Football League

The mandatory constraints considered to schedule the portuguese league were the following:

1. Every team does one, and only, match *per* week;
2. Every team plays against the others only one time in each half of the season;
3. The three biggest clubs (*Benfica*, *Porto* and *Sporting*) can't play against each other on the first two weeks of the tournament;
4. The doubles *Benfica* and *Sporting*, *Porto* and *Boavista*, *Braga* and *Guimarães* and *Marítimo* and *Nacional* can't have home or away matches simultaneously, since they share the same region and a rivalry between them does exist;
5. Games traditionally nominated by *classics* and *derbys* (games between the three biggest clubs) can't be realized in consecutive weeks;
6. *Benfica*, *Porto* and *Sporting* can't play all the two matches against each others in home, at the first half of the season;and,
7. The second half of the season must be a mirror of the first *i.e.*, same match order along the second part with only the game venues exchange: *mirrored tournament*.

Since all of this points were applied as raffle conditionals on the 2014/2015 official season, it was considered mandatory to satisfy all of them in the developed model. The objectives or criteria to be optimized are two, and they can be explained as follows:

- Minimization of the number of home, or away, consecutive games disputed by the same team (applied to all teams, during the season). This event is called a *break*, in international bibliography, as previously introduced; and,
- Minimization of the number of consecutive games between the same smaller team against two of the three biggest ones.

These two criteria will constitute the total objective function of this scheduling model. The objective is to bring more competition to the league by adding fairness and competitiveness to the game with, essentially, the optimization of the number of *breaks* and with concerning smaller teams' reality, by sparing them of playing the hardest games consecutively and avoiding the possibility of loosing many points straight in two followed journeys.

2.2 The Referee Assignment

Nowadays, referees are commonly criticized by players, coaches or even club representatives when their teams loose a match. This kind of controversy creates tensions between referees, clubs and the league organization responsible. The capacity of choosing the right referee to every game isn't an easy task. Many factors depend on that decision and an old question can, now, be raised: should the referees be

assigned to matches by nomination or by lottery? Well, this is a difficult question with an ambiguous answer. Portugal and other major European leagues implement nomination schemes. This article helps on providing methods capable of assigning referees to matches by lottery under certain premisses, and by that it gives some relevance to this possible alternative. The first approach consists in assigning referees, in the beginning of the season, to all of the games, while the second, closer to the portuguese reality, assigns referees on a weekly basis considering their classification along the entire season. Both respect every restrictions and criteria if considered.

For both cases, it will only be considered one match referee, not considering the others that complete a referee's judging team: the forth referee and the two line referees. The constraints considered to model this problematic were the following:

1. Existence of a total of 23 available referees, 9 of which have international category;
2. To every match it will only be assigned one referee;
3. One referee can only judge one match per week;
4. Every referee has to judge a minimum of 8 games in the entire season;
5. In the presence of a classic game (matches between *Benfica, Porto* and *Sporting*), the referee's assignment goes to the international ones;
6. If a referee is assigned to a team in any week, he can only referee a match with that same team three weeks later; and,
7. Every referee of the 23 available is capable of judging any game in the tournament.

Additionally, an eighth restriction will be created and added to the model, in order to avoid the assignment of the same referee to the matches between the same two teams in both rounds of the competition. It is believed that this can avoid possible conflicts and tensions between clubs and the organization that is in charge of this process, before and after those games.

3 Methodology

On this entire study, an half of a season can also be named as a *round* or a *part* and a week may also be called *journey*. The Portuguese league is a competition with two halves (most commonly used in football), but other organisation systems can exist like a single round robin tournament (only one part) or even a round robin tournament with three parts like the Danish one, studied by Rasmussen (2008).

The difference between both parts, in the Portuguese case, is the change of the venues where the matches occur. For example, if in the first journey of the first round, team A plays home against team B, in the first journey of the second round it will be team B that will play on its venue against team A. The order of the games from one part to the other will remain the same.

Note that in England, for instance, a different system is used, which differs from ours in the order of the games played in the second half (reverse of the first). The games disputed in the last journey of the first half will be the same games disputed in the first journey of the second half, and so on...

The portuguese league is constituted by 18 teams. Every team plays in every journey which makes 9 games *per* week, 17 weeks *per* round and a total of 306 games per season. A match is disputed by a pair of teams *(i,j)*, where it is assumed that team i plays *home* and team j plays *away*.

A very useful approach used along this years of investigation about sports scheduling, not only in football, but also in many other different sports, like basketball, (Nemhauser and Trick, 1998), handball, (Larson and Johansson, 2013), rugby, (Lewis and Thompson, 2011), or cricket, (Willis and Terrill, 1994), is the *home-away pattern*. Basically, this term represents, for each team, a vector consisting in H (home) and A (away) characters, of length equal to the number of the weeks of a single part of the competition. When talking about double round robin tournaments, just considering half of the season will be enough, because the other part will be the complementary of the first. A set of home-away patterns is, than, a set of patterns correspondent to all the teams and their games and can be represented, for one team for instance, as follows: $I : \{AHAHAHAHAHAHAHAHA\}$. In this case, certain team I will play *away* in the first week, *home* in the second and, once again, *away* on round 3, and so on through all the necessary journeys to play one game against each opponent. Different occurrences can happen like, for instance : $\{HAAHAH...H\}$, here the phenomena called *break* occurs in weeks two and three and it happens every time a team plays two home, or away, games on two consecutive journeys. So, a complete championship schedule can be represented as a matrix, with A and H characters, with the dimensions n vs *(n-1)*, where n represents the number of competing teams and *(n-1)* the journeys, which is the same of the number of the opponents (or games) for each team.

As previously said, one of the objectives of the scheduling model will be to minimize the number of existing *breaks* and to build a better timetable when compared to the real existing one (season 2014/2015).

After this insights, it is now appropriate to introduce the needed formalism for the creation and development of both integer models.

3.1 The Portuguese Football League Model

The modelling of this problem will only be made for the first part of the season, since the second is a mirror of the first one and can easily be provided when programmed in any suitable software. Obviously this can only be done, because all the conditions will remain the same in both parts. When translating the model to a computational language, it is easy to adjust the rounds of the timetable by adding the number of weeks corresponding to an half and by taking into account that the pairs of games now change from *(i,j)* to *(j,i)*.

SETS

E: set of teams, $E = \{1, ..., n\}, n = 18$
$E1$: subset of E that represents the best teams of the competition, $E1 = \{1, ..., n_1\}$; $n_1 = 3$: *Benfica*, *Porto* and *Sporting*
K: set of weeks, $K = \{1, ..., n-1\}$
K': subset of K, representing the first journeys of the competition (n'), where games between the three biggest clubs can't happen, $K' = \{1, .., n'\}$
C: set of regions that have more than one team. $E(c) \subset E$

For example, $E(A)$ is a subset of E that defines the teams existing in region A. For this case, four regions will be considered (***Lisboa***- *Benfica and Sporting,* ***Porto***- *Boavista and Porto,* ***Madeira***- *Nacional and Maritimo and* ***Minho***- *Braga and Guimarães*).

INDEXES

i- represents the team that plays *home;* j- represents the team that plays *away;* k- represents journeys or weeks.

DECISION VARIABLE
$x_{i,j,k} = 1$, if team i plays against team j on week k and 0 if not.

CONSTRAINTS

1. $\sum_{i \in E}(x_{i,j,k} + x_{j,i,k}) = 1, \forall j \in E, \forall k \in K$ and $i \neq j$
2. $\sum_{k \in K}(x_{i,j,k} + x_{j,i,k}) = 1, \forall (i,j) \in E^2$ and $i \neq j$
3. $\sum_{(i,j) \in E1^2}(x_{i,j,k} + x_{j,i,k}) = 0, \forall k \in K'$ and $i \neq j$
4. $\sum_{(w_1,w_2) \in E^2}(x_{i,w_1,k} - x_{w_2,j,k}) = 0, \forall (i,j) \in E(c)^2, \forall k \in K$ and $i \neq j, j \neq w_2, i \neq w_1$ and $w_2 \neq w_1$
5. $\sum_{(i,j) \in E1^2}(x_{i,j,k} + x_{j,i,k+1}) \leq 1, \forall k \in (1, ..., n-2)$ and $i \neq j$
6. $\sum_{k \in K}\sum_{j \in E1} x_{i,j,k} = 1, \forall i \in E1$ and $i \neq j$
7. $\sum_{k \in K} x_{i,j,k} - \sum_{k \in (K+(n-1))} x_{j,i,k} = 0, \forall (i,j) \in E^2$ and $i \neq j$

The first two restrictions guarantee the basic organizational and structural characteristics of this type of tournaments by making mandatory that each team has only one game *per* week and that every team plays against each other once, in each part of the season. Restriction 3 avoids matches between the (three) biggest clubs on the first n' (=2) journeys of the competition. Restriction 4 provides complementary calendars to the teams from the same geographical region. The w indexes represent other teams but i or j. Next, the restriction 5, avoids the scheduling of classical games on consecutive weeks, while restriction 6 guarantees that top clubs ($E1$) can't play against all the others of the same category home, or away, in an half of the season. Finally, the seventh restriction says that the second half of the season is a mirror of the first, as for the mirrored tournament definition.

OBJECTIVE FUNCTION

As previously said and now reinforced, the two criteria considered for this bi-objective integer model are (I) the *break* minimization and (II) the minimization of the number of consecutive games between the same smaller team against two of the biggest.
Both can be mathematically modelled as follows:

(I) First criteria:

- $min\ [2 \times (\sum_{i \in E}\sum_{k=1}^{K-2}(\sum_{j \in E, j \neq i}(x_{i,j,k} + x_{i,j,k+1}) - 1)^2) + \sum_{i \in E}(\sum_{j \in E, j \neq i}(x_{i,j,(n-1)} + x_{j,i,1}) - 1)^2]$

(II) Second criteria:

- $min\ [2 \times (\sum_{i \in E}(\sum_{j \in E1, j \neq i}\sum_{w \in E1, w \neq j}\sum_{k=1}^{K-2}(x_{i,j,k} \times x_{i,w,k+1} + x_{j,i,k} \times x_{w,i,k+1} + x_{i,j,k} \times x_{w,i,k+1} + x_{j,i,k} \times x_{i,w,k+1}))) + \sum_{i \in E}(\sum_{j \in E1, j \neq i}\sum_{w \in E1, w \neq j}(x_{i,j,(n-1)} \times x_{w,i,1} + x_{j,i,(n-1)} \times x_{i,w,1} + x_{i,j,(n-1)} \times x_{i,w,1} + x_{j,i,(n-1)} \times x_{w,i,1}))]$

Due to the non linearity of both criteria, the function will be treated separately from the rest of the problem. Which means that *CPLEX ilog Studio v.12* was used to obtain as many feasible solutions as possible from the integer modelling of every restriction, and than, those were imported to *Wolfram Mathematica 9.0* where, by the linear definition of both events (criteria) and by the implementation of the objective function on every solution found, the problem was computationally optimized. Than the event occurrences will be counted and the minimized schedules for both criteria will be recorded. After that, the best result, of the top three solutions, will be taken and individually analysed. The results will be shown further on this work. This approach was developed because *CPLEX* is not prepared to work with non linearity and it was necessary to find an alternative way to bypass this issue. Another alternatives would also be interesting to implement and to compare the results, such as the introduction of heuristics. The challenge is, by now, launched for future investigators on this area! For now, to conclude this subsection, what matters is to algebraically finish modelling the problem, and that was successfully achieved with the development and final presentation of the objective function.

3.2 The Referee Assignment Model

Since the scheduling model is considered preliminary to this one, as it will be used as an *input* in order to obtain results, it is assumed that, at this stage, a complete timetable is already provided and the referee assignment must be done for the entire season, based and applied on it.

For this resource assignment problem, two different paths will be explored. First it won't be considered any weekly factor of attribution. Given the full schedule, it will only be assigned one referee to every single match, of the 306 existing games. This approach does not represent how it is done in reality. In Portugal, the referee's nomination is made on a weekly basis and it considers different factors, such as referee's physical and psychological conditions and the previous match classification. Despite that and since the actual process is manually made, the implementation of an automatic assignment base, that satisfies every restrictions, will increase the value of the process, as it will turn it more efficient, effective and, eventually, impartial. To approximate this model to reality, a second approach will be developed and it will already consider the referee's classifications along the competition to influence the assigning process. Usually, this classification is attributed from 1 to 5 to each referee, depending on his performance.

The following model formulation refers to the first approach and it won't consider any objective function on the assignment. The main objective is only to satisfy all the restrictions that are currently used in reality.

SETS

E: set of teams, $E = \{1, ..., n\}, n = 18$
K: set of weeks, $K = \{1, ..., 2(n-1)\}$
A: set of referees, $A = A_{int} \cup A_{nint}$,where A_{int} is the subset of international referees and A_{nint} is the subset of non-international referees
$A_{int} = \{r_1, ..., r_x\}$
$A_{nint} = \{r_{x+1}, ..., r_{xfinal}\}$
J: set of games, $J = J_c \cup J_n$,where J_c is the subset of classical games and J_n is the subset of normal games (non-classical games)

INDEXES

i- index representative of referees; j- index representative of games; e- index representative of teams; k- index representative of journeys.

PARAMETERS

$\alpha_{j,k}$: 1 if game j is played on week k and 0 if not
$\beta_{j,e}$: 1 if in game j plays team e and 0 if not
a_{min}: Minimum number of games that a referee should judge in a full season; will be considered as 8
$a_{i,e}$: Maximum number of games judged by the referee i to team e, in a season
t: Minimum number of weeks that a referee has to be out from judging matches of the same team e; will be considered an interval of 3

DECISION VARIABLE

$x_{i,j} = 1$ if referee i is assigned to match j and 0 if not.

CONSTRAINTS

1. $\sum_{i \in A} x_{i,j} = 1, \forall j \in J$
2. $\sum_{j \in J} x_{i,j}\alpha_{j,k} \leq 1, \forall i \in A$ and $k \in K$
3. $\sum_{k \in K} x_{i,j} \geq a_{min}, \forall i \in A$
4. $\sum_{i \in A_{int}} x_{i,j} = 1, \forall j \in J_c$
5. $\sum_{d=0}^{t-1} \sum_{j \in J} x_{i,j}\beta_{j,e}\alpha_{j,k+d} \leq 1, \forall i \in A, \forall e \in E$ and $[k \leq |k| - c + 1]$
6. $\sum_{k \in K} (x_{i,j}\alpha_{j,k} + x_{i,j}\alpha_{j,k+(n-1)}) = 1, \forall j \in J$ and $i \in A$
7. $\sum_{j \in J} x_{i,j}\beta_{j,e} \leq a_{i,e}, \forall i \in A$ and $\forall e \in E$

The first two restrictions guarantee that to each game it will only be assigned one referee and that a referee will judge at most one game *per* week. The restriction 3 forces the achievement of a minimum number of games by each referee. The restriction 4 assigns to the classical games referees with international statute and imposes no conditions to the other games. Restriction 5 defines that a referee can't be assigned to the same team more than once in t consecutive weeks. Finally, restrictions 6 and 7 state, respectively, that a referee can't judge matches between the same two teams and that should be stipulated a maximum number of games a referee can judge to the same team. This last restriction represents an additional possibility to be considered by the decision maker, already modelled and ready to be computationally implemented, but it won't be applied on this case.

For the second approach a few changes must be done. The adopted procedure will be explained next:

- **Week 1:** The referee assignment for all the games disputed in this journey is based only on the restrictions provided from the previous approach. At this stage, any factor related to the referee performances is considered, since the competition is beginning . It is assumed that every referee starts the season in a good physical and psychological shape and capable of judging every game.
- **Week 2:** Here the objective is to assign the unused referees to games. It's necessary to develop an objective function capable of giving priority to that, which can be given by:
  $max \sum_{i \in A_{free}} \sum_{j \in \{9\times(k-1), 9k\}} x_{i,j}$
  The set of unused referees is defined as A_{free}

and the correspondent games to every journey of the competition are given by $\{9 \times (k-1), 9k\}$, where k represents the respective journey on which its wanted to know the interval of realized games (here, *k=2*). All of them will be numbered from the first to the last game of the competition. Which means that week 1 has games with indexed numbers from 1 to 9, the second week from 10 to 18, and so on...

- **Week 3:** At this stage there are a few referees that will still judge their first match (4 out of the 23, since the sum of the previous disputed games is 18) while the rest will attend games for their second time. The objective will be to assign the unused referees and than complete the other slots with 5 of the rest of them.

- **Weeks 4 to 33:** Among this period the assignment is made in the same way and will take into account all of the already used restrictions, plus the following criteria:

  - Priority to the assignment of the less used referees;
  - Priority to the assignment of referees which had maximum classification on the previous week (equal to five);
  - Priority to avoid the attribution of referees whose classification was lower or equal than two on the previous week.

The objective function can be represented as follows:

$$max\ [0.5 \times (\sum_{i \in A_{free}} \sum_{j \in \{9\times(k-1),9k\}} x_{i,j}) + 2 \times (\sum_{i \in A_{classif_i=5}} \sum_{j \in \{9\times(k-1),9k\}} x_{i,j}) - 10 \times (\sum_{i \in A_{classif_i \leq 2}} \sum_{j \in \{9\times(k-1),9k\}} x_{i,j})]$$

The first two parts of the *max* function represent the first two criteria announced before and contribute positively to the maximization of the objective solution, while the last one has a negative impact on it, since it is as better as it is minimized. That's why it was attributed a relatively high weight (10), when comparing to the other ones (0.5 and 2), because the purpose was to avoid that happening as much as possible. Here, the "A_{free}" set is adapted to the set of referees with less games disputed. The used values for the weights of this weighted-sum were developed jointly with the FPF.

- **Week 34:** In the last week of the championship, since it can be the week of the big decisions of the competition, the objective will be to assign, from the available referees, the ones with higher accumulated classification in the entire season or,in other words, the *best*, available, referees of the season:

$$max \sum_{i \in A_{best}} \sum_{j \in \{9\times(k-1),9k\}} x_{i,j}$$

4 Computational Analysis and Applications

This section will present the results obtained from the computational analysis of the models. It will be divided in three main subsections. The first one will take care of the evaluation of the calendar used by *Liga NOS*. Basically all the restrictions used will be validated in *Mathematica* and the number of occurrences of both criteria ((I) and (II) in the previous chapter) will be counted. With that, will be easy to compare the outcome of our model with the real one currently in use in the portuguese league. The second and third subsections will be, respectively, focused on the evaluation of the first and the second models previously developed, respectively game scheduling and referee assignment.

4.1 The actual Portuguese Football League calendar

The official calendar used in season 2014/2015 was introduced in *Mathematica* with the objective of validating all the restrictions used. The restrictions that *Liga NOS* said that were considered and satisfied, were, in fact, present and satisfied. When the objective function was applied on it, the outcome results, for both criteria, were the following:

Table 1: Results of both criteria in 2014/2015 official schedule

Criteria I	**Criteria II**
48	13

Which means that, this timetable has 48 *breaks* and 13 occurrences of smaller teams facing, in consecutive weeks, two of the three biggest clubs.

4.2 First model results: The Portuguese Football League

As said before, since non linearity on the objective function exists, it was necessary to outline this situation. Since it is relatively accessible to define in software *Mathematica* the event "*break*" and the event "*small team playing two consecutive games against a bigger team*", they will be defined and through the implementation of a *counter*, their occurrences will be recorded. This process will occur in every solution from the total of the 11000 feasible solutions found on CPLEX. The objective is to minimize both number of occurrences, by finding the schedule with the lowest values.

A *break* has two possible scenarios (two *home* or two *away* games) and it can be defined as follows:

- $\sum_{(j,w)\in E^2} \sum_{k\in(K-2)} (x_{i,j,k} + x_{i,w,k+1}) = 2, \forall i \in E, w \neq j, i \neq w \text{ and } i \neq j$
- $\sum_{(j,w)\in E^2} \sum_{k\in(K-2)} (x_{j,i,k} + x_{w,i,k+1}) = 2, \forall i \in E, w \neq j, i \neq w \text{ and } i \neq j$

These expressions are constructed to make sense for one half of the season. For the entire championship the results

must be doubled, since the existing games in the first half are the same than those existing in the second. A particular case will rise and will deserve special attention. It is the case of the connection of the transient journeys, which means the last journey of the first half and the first journey of the second, which is not directly accounted by this system. As an example, the *home breaks* on those journeys can be defined as in the following expression:

$$\sum_{(j,w)\in E^2} (x_{i,j,(n-1)} + x_{w,i,1}) = 2, \forall i \in E, w \neq j \text{ and } i \neq j$$

The first journey of the second half is structured by changing the venues of the games of the first journey of the first half. A similar expression for *away* games can be developed.

The definition of *consecutive games between a team against two from the three biggest* can be constructed as follows.

- $\sum_{(j,w)\in E1^2} \sum_{k\in(K-2)} (x_{i,j,k} + x_{i,w,k+1}) = 2, \forall i \in E, w \neq j, i \neq w \text{ and } i \neq j$
- $\sum_{(j,w)\in E1^2} \sum_{k\in(K-2)} (x_{j,i,k} + x_{w,i,k+1}) = 2, \forall i \in E, w \neq j, i \neq w \text{ and } i \neq j$
- $\sum_{(j,w)\in E1^2} \sum_{k\in(K-2)} (x_{i,j,k} + x_{w,i,k+1}) = 2, \forall i \in E, w \neq j, i \neq w \text{ and } i \neq j$
- $\sum_{(j,w)\in E1^2} \sum_{k\in(K-2)} (x_{j,i,k} + x_{i,w,k+1}) = 2, \forall i \in E, w \neq j, i \neq w \text{ and } i \neq j$

It is important to say that, in this situation it is mandatory to define four cases since there is no restrictions about a team playing all games in *home*, *away* or mixed (*home-away or away-home*).

The expressions relative to the transient half journeys can be developed with the same approach used in the last case.

The top ten best calendar results, for both criteria, obtained using both software are shown in **Table 2**.

Table 2: Top ten best results found

Criteria I	Criteria II
43	2
43	2
43	4
43	4
43	4
43	4
43	4
43	4
43	4
43	4

In the 11000 solutions, at least ten of them are considerably better than the actual, for those two criteria. The best calendar found has 43 *break* occurrences and only 2 situations representing criteria II. A portion of the best calendar output will be shown further on, together with the referee's model results.

4.3 Second model results: The Referee Assignment

The referee assignment can be done in two ways. The first one will assign, at moment zero of the competition, all referees to all games, respecting the restrictions considered. In reality the assignment is made on a weekly basis and considers many factors as the match importance, the referee physical and psychological conditions and referee's classifications on their games. The second approach will consider this last point. The totality of these factors won't be considered on this study due to their unpredictability. A future ambition will be to integrate all of them in a future model.

4.3.1 First Approach

The results for the first model under the above mentioned first approach are shown in **Appendix A**. Since the outcome for the entire season would be dense, only the results for the first half of the best calendar found (see **Table 2**) and the respective referee assignment are shown. Despite the fact that the referee's allocation is computationally well done with satisfiable outcomes, it can be clearly seen that, in some moments, the assignment doesn't reflect a realistic logic. Nevertheless this approach can be useful, since its being created in seconds a full assignment, but it isn't realist. It could bring transparency to the process, as it is fair and exempt of "external" factors; additionally it could turn the process more efficient, allowing the referees allocation automatically. The weakness of this approach will be controlled by the introduction of some real criteria, as you will see further on this work.

Note that, as opposed to the clubs, which have the real names, the referees don't have. This is basically because it is assumed that every referee available for the first professional football league has enough category to judge all the existent games. The international referees are represented with numbers from 1 to 9 and the non international ones were considered the rest of them.

4.3.2 Second Approach

The weekly referee's classification was simulated in the same software as the assignment models were developed, *CPLEX*. The generated numbers followed a *uniform distribution* from 1 to 5. The results from this version of the model are presented in **Appendix B**.

To finish this section, all the results provided on this paper were influenced by the conditions of the only available *hardware* characteristics. Throughout the work, it was used an *Intel iCORE 7* processor with *4GBs* of memory *RAM* and *2.4GHz* of processor speed. Obviously, the better the characteristics and other important computational features, the better will be the computer performance and the processing time as well as the set of the outcome results.

5 Conclusions

The construction of a schedule that respects all the considered restrictions and optimizes certain defined criteria, can be done in a total of 3 hours and 9 minutes. This processing time took 9 minutes in *CPLEX*, from where were obtained all the feasible solutions of the scheduling problem and 3 hours on *Mathematica*, where the problem was optimized. Related to the referee assignment problem, the first adopted approach only took 30 seconds to generate results, while the second took 4 seconds *per* week, which can be computed in 136 seconds for the total 34 weeks, approximately 2 minutes of processing time.

It is never too much to reinforce the assumptions that were made during this entire work. First, it was considered that all the referees are available to be assigned during the competition, (no injuries nor clearances were considered). Every week, the ones not assigned for the first football league games are assumed to referee the second league. The weekly approach considers referee's classification for the second football division even without caring about the scheduling of this league. Only cares about the fact that in every week 12 more referees must be assigned and classified. Then, no international competitions were considered nor national cups, on both models. All of this factors can be considered as suggestions and also as possible alternatives that can be explored in future investigation on this area. As much they can be included, the closer to reality the models will be. Despite that, it is believed that the contribution of this work in science was positive, since it explores, in a first phase, issues and a case that were not considered by investigators until this present date. It is believed, from this day on, that Portugal will be represented in the existent worldwide literature about Operations Research in sports. Plus, it is valid to state that two interesting and efficient tools were here constructed in a way that can be used by both organizations that are in charge of games and referees in Portugal (*Liga de Clubes NOS* and FPF, respectively), since all of the scheduling characteristics and the most important points of the referee's assignment are, in here, taken into account, preserving the reality of this case study.

Finally, I hope that this work can provide an important contribute to Operations Research in schools. It is, for sure, one of the greatest ambitions of this entire project. To incentive professors to use more frequently sports as an example to teach Operations Research. Sports can be a fun, appealing and efficient way for stimulating the students attention and productivity, since everyone knows more or less about a sport, or another, to be capable of understanding and connect the main problems intrinsic to the Operations Research academical teaching scope.

5.1 Acknowledgements

First of all, I would like to thank to my family and friends for all the support provided. Then, a special acknowledgement will go to *professor José Rui Figueira* for all the given advices through these entire journey. Additionally, to *Dr. Rui Chumbita Nunes*, who gently introduced me to FPF and *Liga de Clubes Portugal*. Speaking in them, another special acknowledgement must be done to the vice-president of the FPF, *Dr. Rui Manhoso*, to the vice-president of the Referee Council of the the same institution, *Dr. Antonino da Silva* and to the director of *Liga de Clubes Portugal*, *Dra. Andreia Couto*. They all had an important role on this investigation and without all of them the result of this work would never be the same.

References

Alarcón, F., Durán, G., and Guajardo, M. (2012). Referee assignment in the chilean football league using integer programming an patterns. 21(3):415–438.

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

Della Croce, F. and Oliveri, D. (2006). Scheduling the Italian Football League: An ILP-based approach. *Computers and Operations Research*, 33(7):1963–1974.

Dur, G., Guajardo, M., Miranda, J., and Saur, D. (2005). Scheduling the Chilean Soccer League by Integer Programming. 37(6):539–552.

Fiallos, J., Pérez, J., Sabillón, F., and Licona, M. (2010). Scheduling Soccer League of Honduras using Integer Programming. Institute of Industrial Engineers-Publisher.

Flatberg, T. and Nilssen, E. (2009). Scheduling the topmost football leagues of Norway. In *23rd European Conference on Operational Research Book of Abstracts*, page 240.

Kendall, G., Knust, S., Ribeiro, C. C., and Urrutia, S. (2010). Scheduling in sports: An annotated bibliography. *Computers and Operations Research*, 37(1):1–19.

Larson, J. and Johansson, M. (2013). Constructing Schedules for Sports Leagues with Divisional and Round-robin Tournaments. *Journalof Quantit. Analysis in Sport*, 10(2):119–129.

Lewis, R. and Thompson, J. (2011). On the application of graph colouring techniques in round-robin sports scheduling. *Computers and Operations Research*, 38(1):190–204.

Nemhauser, G. L. and Trick, M. a. (1998). Scheduling A Major College Basketball Conference. *Operations Research*, 46(1):1–8.

Rasmussen, R. V. (2008). Scheduling a triple round robin tournament for the best Danish soccer league. *European Journal of Operational Research*, 185(2):795–810.

Recalde, D., Torres, R., and Vaca, P. (2013). Scheduling the professional Ecuadorian football league by integer programming. *Computers and Operations Research*, 40(10):2478–2484.

Ribeiro, C. C. and Urrutia, S. (2014). Scheduling the Brazilian Soccer Tournament : Solution Approach and Practice. 42(3):1–16.

Willis, R. and Terrill, B. (1994). Scheduling the Australian State Cricket using Simulated Annealing. *The journal of the Operational Research Society*, 45(3):276–280.

Yavuz, M., Inan, U. H., and Filali, A. (2008). Fair referee assignments for professional football leagues. *Computers and Operations Research*, 35(9):2937–2951.

6 Appendix

6.1 Appendix A

Table 3: First half of the best calendar found and referee assignment by the first approach

Journey	*Home team*	*Away team*	*Match*	*Referee*
1	Boavista	Arouca	1	3
	Estoril	Penafiel	2	4
	GilVicente	FCPorto	3	21
	Maritimo	Belenenses	4	20
	PFerreira	Moreirense	5	11
	RioAve	SCBraga	6	17
	Sporting	Nacional	7	5
	Guimaraes	Academica	8	9
	Setubal	Benfica	9	18
2	Academica	Boavista	10	7
	Belenenses	Guimaraes	11	17
	FCPorto	Setubal	12	5
	Moreirense	RioAve	13	2
	Nacional	Arouca	14	15
	PFerreira	Benfica	15	4
	Penafiel	Maritimo	16	18
	SCBraga	GilVicente	17	13
	Sporting	Estoril	18	22
3	Arouca	Estoril	19	5
	Benfica	Nacional	20	1
	Boavista	Belenenses	21	14
	GilVicente	Sporting	22	6
	Maritimo	Academica	23	4
	Penafiel	PFerreira	24	9
	RioAve	FCPorto	25	19
	Guimaraes	Moreirense	26	21
	Setubal	SCBraga	27	22
4	Academica	Estoril	28	1
	Belenenses	Penafiel	29	5
	Benfica	Guimaraes	30	12
	FCPorto	Sporting	31	7
	GilVicente	Arouca	32	23
	Moreirense	Maritimo	33	13
	Nacional	RioAve	34	22
	PFerreira	Setubal	35	3
	SCBraga	Boavista	36	18
5	Arouca	RioAve	37	20
	Belenenses	SCBraga	38	21
	Boavista	Benfica	39	11
	Estoril	GilVicente	40	2
	Nacional	Moreirense	41	17
	Penafiel	FCPorto	42	15
	Sporting	Academica	43	12
	Guimaraes	PFerreira	44	8
	Setubal	Maritimo	45	19
6	Academica	Arouca	46	5
	Benfica	RioAve	47	3
	FCPorto	Belenenses	48	4
	GilVicente	Nacional	49	13
	Maritimo	Guimaraes	50	22
	Moreirense	Estoril	51	6
	PFerreira	Boavista	52	1
	Penafiel	Setubal	53	21
	SCBraga	Sporting	54	10
7	Arouca	FCPorto	55	11
	Belenenses	Academica	56	6
	Boavista	Maritimo	57	3
	GilVicente	Guimaraes	58	23
	Nacional	Penafiel	59	14
	PFerreira	Estoril	60	4
	SCBraga	Benfica	61	12
	Sporting	Moreirense	62	16
	Setubal	RioAve	63	2
8	Belenenses	Sporting	64	1
	Benfica	FCPorto	65	8
	Boavista	Nacional	66	6
	Estoril	SCBraga	67	11
	Maritimo	PFerreira	68	16
	Moreirense	Arouca	69	3
	Penafiel	Academica	70	9
	RioAve	GilVicente	71	7
	Guimaraes	Setubal	72	17
9	Academica	RioAve	73	19
	Arouca	Sporting	74	2
	Benfica	Penafiel	75	16
	Estoril	Belenenses	76	22
	FCPorto	Nacional	77	9
	GilVicente	PFerreira	78	10
	Maritimo	SCBraga	79	17
	Guimaraes	Boavista	80	15
	Setubal	Moreirense	81	18
10	Belenenses	Arouca	82	19
	Boavista	Setubal	83	22
	Moreirense	GilVicente	84	21
	Nacional	Maritimo	85	3
	PFerreira	Academica	86	12
	Penafiel	Guimaraes	87	7
	RioAve	Estoril	88	5
	SCBraga	FCPorto	89	1
	Sporting	Benfica	90	4
11	Academica	Setubal	91	20
	Arouca	Benfica	92	5
	Belenenses	Nacional	93	14
	Estoril	Boavista	94	16
	FCPorto	Guimaraes	95	21
	GilVicente	Penafiel	96	1
	Maritimo	RioAve	97	13
	SCBraga	Moreirense	98	6
	Sporting	PFerreira	99	17
12	Benfica	Moreirense	100	20
	FCPorto	Estoril	101	8
	GilVicente	Academica	102	4
	Maritimo	Sporting	103	15
	PFerreira	Belenenses	104	7
	Penafiel	SCBraga	105	11
	RioAve	Boavista	106	2
	Guimaraes	Nacional	107	13
	Setubal	Arouca	108	17
13	Academica	SCBraga	109	9
	Belenenses	RioAve	110	11
	Benfica	Maritimo	111	14
	Boavista	GilVicente	112	13
	Estoril	Setubal	113	4
	Moreirense	FCPorto	114	5
	Nacional	PFerreira	115	6
	Penafiel	Arouca	116	21
	Guimaraes	Sporting	117	19
14	Academica	Benfica	118	1
	Belenenses	GilVicente	119	5
	Boavista	Moreirense	120	14
	Maritimo	Estoril	121	12
	PFerreira	FCPorto	122	16
	RioAve	Guimaraes	123	20
	SCBraga	Arouca	124	19
	Sporting	Penafiel	125	18
	Setubal	Nacional	126	21
15	Arouca	Guimaraes	127	17
	Benfica	Estoril	128	13
	FCPorto	Maritimo	129	20
	GilVicente	Setubal	130	22
	Moreirense	Belenenses	131	2
	Nacional	Academica	132	14
	Penafiel	Boavista	133	23
	RioAve	Sporting	134	1
	SCBraga	PFerreira	135	8
16	Academica	Moreirense	136	16
	Belenenses	Benfica	137	8
	FCPorto	Boavista	138	11
	GilVicente	Maritimo	139	18
	Nacional	SCBraga	140	13
	PFerreira	Arouca	141	15
	Penafiel	RioAve	142	17
	Sporting	Setubal	143	6
	Guimaraes	Estoril	144	4
17	Academica	FCPorto	145	19
	Belenenses	Setubal	146	15
	Benfica	GilVicente	147	3
	Boavista	Sporting	148	5
	Estoril	Nacional	149	23
	Maritimo	Arouca	150	13
	Moreirense	Penafiel	151	10
	RioAve	PFerreira	152	21
	SCBraga	Guimaraes	153	22

6.2 Appendix B

Table 4: First half of the best calendar found and referee assignment by the second approach

Journey	*Match*	*Referee*	*Classification*
	1	3	4
	2	4	4
	3	21	1
	4	20	3
	5	11	1
	6	17	1
	7	5	5
1	8	9	1
	9	18	5
	10	23	3
	11	8	4
	12	10	5
	13	15	1
	14	12	5
	15	1	2
	16	19	1
2	17	2	4
	18	14	4
	19	2	2
	20	16	2
	21	6	4
	22	3	2
	23	1	5
	24	7	2
	25	22	3
3	26	13	5
	27	4	2
	28	8	3
	29	1	4
	30	5	1
	31	9	5
	32	13	1
	33	11	4
	34	6	4
4	35	22	1
	36	21	3
	37	15	2
	38	18	1
	39	23	5
	40	12	4
	41	14	1
	42	17	1
	43	16	3
5	44	11	5
	45	9	1
	46	10	3
	47	16	5
	48	12	1
	49	23	3
	50	20	4
	51	3	1
	52	2	4
6	53	6	5
	54	11	5
	55	16	1
	56	6	4
	57	19	1
	58	14	5
	59	10	3
	60	20	4
	61	1	5
7	62	23	2
	63	11	5
	64	14	4
	65	8	5
	66	11	1
	67	5	5
	68	12	5
	69	4	1
	70	19	4
8	71	1	2
	72	13	4
	73	5	1
	74	12	2
	75	2	2
	76	18	2
	77	19	2
	78	17	3
	79	8	1
9	80	3	1
	81	14	1
	82	13	1
	83	6	5
	84	15	1
	85	22	1
	86	9	3
	87	17	1
	88	4	3
10	89	10	4
	90	7	5
	91	7	4
	92	1	3
	93	8	2
	94	16	1
	95	6	4
	96	3	5
	97	11	5
11	98	12	2
	99	4	4
	100	18	2
	101	3	3
	102	11	2
	103	15	2
	104	17	2
	105	22	1
	106	7	1
12	107	14	1
	108	9	1
	109	16	1
	110	23	3
	111	8	3
	112	21	3
	113	19	2
	114	1	5
	115	6	4
13	116	20	5
	117	3	5
	118	7	3
	119	10	4
	120	3	4
	121	22	2
	122	20	5
	123	1	2
	124	21	1
14	125	13	3
	126	18	5
	127	11	2
	128	23	1
	129	18	2
	130	14	2
	131	13	2
	132	20	3
	133	16	2
15	134	17	4
	135	2	5
	136	8	3
	137	15	2
	138	2	1
	139	21	2
	140	19	5
	141	9	5
	142	4	2
16	143	12	1
	144	17	4
	145	9	4
	146	16	1
	147	20	2
	148	19	3
	149	7	2
	150	23	3
	151	3	4
17	152	22	1
	153	5	1