ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Briskorn, Dirk; Drexl, Andreas

Working Paper — Digitized Version

Integer programming models for round robin tournaments

Manuskripte aus den Instituten für Betriebswirtschaftslehre der Universität Kiel, No. 613

Provided in Cooperation with:
Christian-Albrechts-University of Kiel, Institute of Business Administration

Suggested Citation: Briskorn, Dirk; Drexl, Andreas (2006) : Integer programming models for round robin tournaments, Manuskripte aus den Instituten für Betriebswirtschaftslehre der Universität Kiel, No. 613, Universität Kiel, Institut für Betriebswirtschaftslehre, Kiel

This Version is available at:
https://hdl.handle.net/10419/147668

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

Manuskripte
aus den
Instituten für Betriebswirtschaftslehre
der Universität Kiel

No. 613

Integer programming models for round robin tournaments

Dirk Briskorn, Andreas Drexl

December 2006

Dirk Briskorn, Andreas Drexl
Christian-Albrechts-Universität zu Kiel,
Institut für Betriebswirtschaftslehre,
Olshausenstr. 40, 24098 Kiel, Germany,
http://www.bwl.uni-kiel.de/bwlinstitute/Prod
briskorn@bwl.uni-kiel.de,andreas.drexl@bwl.uni-kiel.de

Abstract

Generally speaking in sports competition a number of teams play against each other over a period of time according to a certain scheme. The round robin scheme is very popular in many team sports like soccer and basketball. In this paper we study several basic variants of round robin tournaments. Additionally, we examine real world issues that have to be considered when constructing a sports league schedule. We consider constraints imposed by third parties such as security aspects and legal requirements. Moreover, we analyse constraints aiming at attractive and fair sports league schedules. Throughout this paper integer programming models are the means to formally define what the particular problem under consideration is all about. Moreover, the standard solver Cplex is used in order to compute optimal solutions for small- to medium-size instances. Computational results indicate which particular constraint to relax in order to come up with acceptable schedules in reasonable time.

Keywords: Round robin tournaments, integer programming models, real world constraints, computational results

1 Introduction

Round robin tournament (RRT) covers a huge variety of different problems arising in practice. The focus in this paper is on RRTs, where scheduling is temporally constrained which means that a minimum number of periods are given the matches have to be scheduled in. We consider a set T of n teams. If n is odd we can add a dummy team and, hence, we can assume, that n is even without loss of generality.

In a r-RRT each team plays r times against each other, either at home or away. Each team has to play at least $\lfloor \frac{r}{2} \rfloor$ times at home against each other team. Obviously, this implies that no team can play more than $\lceil \frac{r}{2} \rceil$ times at home and the resulting schedule is balanced with respect to the venues of the matches. Furthermore, a team $i \in T$ has to play exactly once in each period and, hence, we have a set P of $r(n-1)$ periods altogether. For the sake of shortness periods are also called matchdays, abbreviated as MD, in the following.

In particular we consider the special cases $r = 1$ and $r = 2$. Real world examples for both cases are: Single RRTs ($r = 1$) often are carried out in major sports events such as FIFA soccer world cup before the play off rounds. Double RRTs ($r = 2$) are predominant in premium class soccer leagues. Illustrative examples for single RRTs and double RRTs are given in tables 1 and 2, respectively.

period	1	2	3	4	5
match 1	1-2	5-6	3-4	4-5	5-1
match 2	5-3	1-4	2-5	3-1	4-2
match 3	4-6	2-3	1-6	2-6	3-6

Table 1: Single RRT for $n = 6$

Models for sports league scheduling have been the topic of extensive research. A whole stream of papers is based on the analogy between sports league scheduling and edge coloring of complete graphs. Examples are de Werra [6, 7, 8, 9, 10], de Werra et al. [11], and Drexl and

period	1	2	3	4	5	6	7	8	9	10
match 1	1-2	6-5	5-6	4-5	5-1	5-4	3-4	2-1	1-5	4-3
match 2	5-3	4-1	1-4	3-1	4-2	1-3	2-5	3-5	2-4	5-2
match 3	4-6	3-2	2-3	2-6	3-6	6-2	1-6	6-4	6-3	6-1

Table 2: Double RRT for $n = 6$

Knust [12]. Brucker and Knust [5] and Drexl and Knust [12] analyze the relationship between sports league scheduling and multi-mode resource constrained project scheduling. Bartsch [1], Bartsch et al. [2], and Schreuder [13, 14] examine particular formulations.

The remainder of this paper is organized as follows. In section 2 we address structural properties of sports league schedules. Additionally, we provide integer programming models which will serve as the base for further constraints to be analyzed. Constraints proposed in section 4 are given by the tournaments organizers while those introduced in section 3 are influenced by outer league parties. The results of an extensive computational study are provided in section 5 and, finally, section 6 contains a summary and an outlook to future work.

2 Basic Models

2.1 Single Round Robin Tournament

Single RRTs are RRTs for the special case of $r = 1$ where each team $i \in T$ meets each other team $j \in T, j \neq i$, exactly once. Also each team $i \in T$ plays at each MD $p \in P$ exactly once. The cardinality of the set P is equal to $n - 1$. We associate "cost" $c_{i,j,p}$ with each match of team $i \in T$ playing at home against team $j \in T, j \neq i$, at MD $p \in P$. Next, we introduce an integer program representing the goal of finding the minimum cost single RRT.

Single RRT-IP

$$\min \sum_{i \in T} \sum_{j \in T \setminus \{i\}} \sum_{p \in P} c_{i,j,p} x_{i,j,p} \tag{1}$$

$$\text{s.t.} \sum_{p \in P} (x_{i,j,p} + x_{j,i,p}) = 1 \qquad \forall\, i, j \in T, i < j \tag{2}$$

$$\sum_{j \in T \setminus \{i\}} (x_{i,j,p} + x_{j,i,p}) = 1 \qquad \forall\, i \in T, p \in P \tag{3}$$

$$x_{i,j,p} \in \{0, 1\} \qquad \forall\, i, j \in T, i \neq j, p \in P \tag{4}$$

Binary variable $x_{i,j,p}$ is equal to 1 if and only if team i plays at home against team j at MD p. The objective function (1) represents the goal of cost minimization. Constraint (2) assures that each team i meets each other team j in exactly one out of the set P of MDs. Although we consider the venue of the match between i and j it does not matter at which venue it is carried out as far as feasibility is concerned. According to (3) each team i plays exactly once per period. Constraints (2) are called 1-factor constraints and constraints (3) are denoted all-different constraints in the literature; see for example Trick [15].

The Single RRT-IP can be used to cover r-RRTs with mirrored rounds where $r \in \mathbb{N}$, $r > 1$. We say that a RRT's rounds are mirrored if a match of i against j at i's home is scheduled at MD $p+n-1$ if and only if a match of i against j at j's home is scheduled at MD p with $p \leq (r-1)(n-1)$. Given a mirrored r-RRT with cost $c'_{i,j,p}$ we can construct a Single RRT-IP by letting

$$c_{i,j,p} = \sum_{r'=1}^{\lceil \frac{r}{2} \rceil} c'_{i,j,(p+2(r'-1)(n-1))} + \sum_{r'=1}^{\lfloor \frac{r}{2} \rfloor} c'_{j,i,(p-n+1+2r'(n-1))} \quad \forall i,j \in T, p \leq n-1.$$

2.2 Multiple Round Robin Tournament

In practice r-RRTs are not always restricted to mirrored rounds, a case which can be covered by the Single RRT-IP as outline above. These r-RRTs are studied in what follows.

A round-based r-RRT consists of $r(n-1)$ MDs partitioned into r rounds where each round forms a single RRT. The subset of periods round r' is arranged in is denoted by $P_{r'}$ with $|P_{r'}| = n-1$ in the following. A match between i against j is carried out at i's home in round $r' < r$ if and only if the corresponding match in round $r'+1$ takes place at j's home. Given cost $c'_{i,j,p}$ for each match the minimum cost round-based multiple RRT is displayed in equations (5) to (9).

Round-based multiple RRT-IP

$$\min \sum_{i \in T} \sum_{j \in T \setminus \{i\}} \sum_{r' \leq r} \sum_{p \in P_{r'}} c'_{i,j,p} x_{i,j,p} \tag{5}$$

$$\text{s.t.} \sum_{p \in P_1} (x_{i,j,p} + x_{j,i,p}) = 1 \quad \forall\, i,j \in T, i < j \tag{6}$$

$$\sum_{j \in T \setminus \{i\}} (x_{i,j,p} + x_{j,i,p}) = 1 \quad \forall\, i \in T, p \in P_1 \tag{7}$$

$$\sum_{p \in P_{r'}} x_{i,j,p} - \sum_{p \in P_{r'+1}} x_{j,i,p} = 0 \quad \forall\, r' < r, i,j \in T \tag{8}$$

$$x_{i,j,p} \in \{0,1\} \quad \forall\, i,j \in T, i \neq j, p \in (P_1 \cup P_2) \tag{9}$$

Constraints (6) and (7) form a single RRT in round 1. Equation (8) forces a single RRT in each round and requires the venue of a specific pair of teams i,j to alternate between the two opponents' stadiums.

A general r-RRT covers a set P of $r(n-1)$ MDs and hosts r matches between i and j no less than $\lfloor \frac{r}{2} \rfloor$ of which are hold at each opponents venue. Given costs $c''_{i,j,p}$ the integer program is provided in equations (10) to (14).

Again, (10) aims at minimum cost. The number of matches between two teams i,j have to be equal to r due to (11). The number of those matches hold at i's home must not differ by more than 1 from the number of matches carried out at j's home according to (12).

Structural properties of RRT problems are studied in Briskorn et al. [4]. In particular it is proven that each of the models defined above belongs to the class of NP-hard optimization

Multiple RRT-IP

$$\min \sum_{i \in T} \sum_{j \in T \setminus \{i\}} \sum_{p \in P} c''_{i,j,p} x_{i,j,p} \tag{10}$$

$$\begin{aligned}
\text{s.t.} \sum_{p \in P} (x_{i,j,p} + x_{j,i,p}) &= r && \forall\, i,j \in T, i < j && (11)\\
\sum_{p \in P} x_{i,j,p} &\geq \left\lfloor \frac{r}{2} \right\rfloor && \forall\, i,j \in T, i < j && (12)\\
\sum_{j \in T \setminus \{i\}} (x_{i,j,p} + x_{j,i,p}) &= 1 && \forall\, i \in T, p \in P && (13)\\
x_{i,j,p} &\in \{0,1\} && \forall\, i,j \in T, i \neq j, p \in P && (14)
\end{aligned}$$

problems. In doing so the case where the costs are restricted to $\{0,1\}$ plays a dominant role. Subsequently we summarize some aspects of what the "cost" might cover.

2.3 Objective Function Coefficients

So far we used the abstract term "cost". In order to emphasize the practical relevance of the concept of cost minimization in this application domain we summarize some aspects of what the "cost" $c_{i,j,p}$ of real-world tournaments might cover: (i) Teams usually have preferences for playing at home at certain MDs, a fact which can easily be expressed through $c_{i,j,p}$. (ii) Since a major objective of the organizers of a tournament is to maximize attendance we can represent the economic value of the estimated attendance by $c_{i,j,p}$ and tackle the maximization version. (iii) Often, a stadium is owned by some public agency and teams do have to pay a fee for each match taking place in that particular stadium. This fee can be represented by $c_{i,j,p}$. (iv) In terms of more complex models the single RRT might be used as a subproblem, e.g., within a Lagrangean relaxation or a column generation framework. Then, $c_{i,j,p}$ is used to cover dual information also. (v) A special case of the single RRT arises when the costs are restricted to $\{0,1\}$. Then $c_{i,j,p} = 1$ denotes that team i cannot play team j in team i's home venue in period p, whereas $c_{i,j,p} = 0$ denotes that this is possible. What we are interested in is to determine whether a feasible schedule, that is, a zero-cost schedule, exists or not.

In Briskorn et al. [4] the case when the costs are restricted to $\{0,1\}$ is referred to as availability constrained single RRT, which is of relevance because of two reasons: (i) First, stadiums may not always be available (leading to $c_{i,j,p} = 1$ for all j if team i's home venue is not available in period p), or team i can, in some period p, only play in stadiums not too far away from its home base. Moreover, if team i does not want to play away in period p then $c_{j,i,p}$ is set to 1 for all j. (ii) Second and more important, the availability constrained single RRT serves as a means to formally state the complexity status of minimum cost RRT problems.

3 Externally Given Constraints

In this section, we consider several requirements of practical relevance usually given by outer league parties. These constraints cannot be influenced by the league's organizers. We represent

them by means of IP constraints. We choose the Single RRT-IP introduced in section 2 as structural core for what follows. This is justified by the fact that many professional soccer, ice hockey, and handball leagues compete according to mirrored double RRTs which is shown to be solvable via single RRT in section 2. Consequently, we project requirements resulting from the second round of a mirrored double RRT to the unique round of single RRT. Furthermore, there are many mega events, e.g. soccer world cup, having several single RRTs followed by a play-off phase. However, adaptation to other RRTs presented in sections 2 is straightforward.

3.1 Forbidden Matches

As outlined in section 1 matches are carried out in one of both teams' stadiums. Although a particular stadium is associated to each team the stadium might not be owned by the club. Often the corresponding city or some public agency is owner or co-owner of the venue. Hence, stadiums generally are used for hosting other events such as pop concerts, too. When scheduling a sports league we can not take stadium availability for granted but have to take care of occupied stadiums. This means, we have to arrange a match at the opponent's venue for team i if i's stadium is not available. The other way round, i has to play at home in period p if its stadium is not available in the corresponding period of the second round in a mirrored double RRT. We introduce parameter $\pi_{i,j,p}$ indicating whether a match (i,j,p) is possible ($\pi_{i,j,p} = 1$) or not ($\pi_{i,j,p} = 0$) in period p and employ it in $n(n-1)^2$ constraints according to (15).

$$x_{i,j,p} \leq \pi_{i,j,p} \quad \forall i,j \in T, i \neq j, p \in P \tag{15}$$

Trivially, constraint (15) forces $x_{i,j,p}$ to be equal to 0 if match (i,j,p) is not possible. If i's stadium is not available in period p we set $\pi_{i,j,p} = 0$ for all $j \in T \setminus \{i\}$. Consequently, if i's stadium is not available in period p of the second round of a mirrored double RRT we set $\pi_{j,i,p} = 0$ for all $j \in T \setminus \{i\}$.

Furthermore, we can fix i's stadium as the venue and p as the period where a match of team i at home against team j is carried out by setting $\pi_{i,j',p'} = 0$ for all (i,j',p), (i,j,p'), (j'',i,p), and (j,i,p'') with $j'' \in T$, $j' \in T \setminus \{i\}$, $p'' \in P$, and $p' \in P \setminus \{p\}$. Analogously, we can force matches whose teams are supposed to be among the best ones to be carried out in a "later" period in order to create a thrilling finale phase of tournament.

3.2 Regions' Capacity

In real world leagues there are regions in which more than one team is located. Consider, for instance, European professional soccer leagues where some capitals host more than one team. A prominent example is London where no less than 13 professional teams are located. Five of them played in the Premier League, England's first soccer league, in season 2005/2006. Even if those teams have a stadium on their own the infrastructure of the region might be overloaded if too many teams play at home in period p. For example, traffic jams and overcrowded public transportation systems resulting from fans heading to the stadiums at the same time must be avoided. Furthermore, the capacity of security staff and of firemen needed in case of emergency is limited.

In order to model the limitation of the number of matches carried out in a specific region in each period we introduce the set R of regions with each $R' \in R$ being a subset of T. In most real world examples $|R' \cap R''| = 0$ holds for all $R', R'' \in R, R' \neq R''$. However, there is no formal need for R either to form a disjunctive or a complete partition of T. We associate a maximum number of matches $C_{R'}$ per period with each $R' \in R$. Parameter $C_{R'}$ has to be customized for each region taking into account all individual aspects such as region's expansion, road network and railway system.

$$\sum_{i \in R'} \sum_{j \in T \setminus \{i\}} x_{i,j,p} \leq C_{R'} \quad \forall R' \in R, p \in P \tag{16}$$

$$\sum_{i \in R'} \sum_{j \in T \setminus \{i\}} x_{j,i,p} \leq C_{R'} \quad \forall R' \in R, p \in P \tag{17}$$

When scheduling a single RRT only $|R|(n-1)$ constraints according to (16) are employed. The number of matches in each region R' is forced to be no more than $C_{R'}$ in each period. In case of a mirrored double RRT we add constraint (17) in order to limit the number of matches in region R' in period p of round 2 by limiting the number of teams of region R' playing away in period p of the first round. Note that, obviously, $C_{R'}$ can not be set to a value less than $\left\lceil \frac{|R'|}{2} \right\rceil$ for a mirrored double RRT since the left hand sides of restrictions (16) and (17) sum up to $|R'|$.

3.3 Highly Attended Matches

Broadcasting stations grew more and more important for leagues and teams. This could be clearly observed when Leo Kirch, owner of german broadcasting station "Premiere", suffered insolvency and as a result many german soccer clubs had a high budget deficit.

A match's attendance depends among others on the two teams competing. Based on, e.g., the score in the previous competition we can identify matches having high attractiveness for spectators and, hence, for broadcasting stations, too. Broadcasting stations are interested in presenting as many attractive matches as possible. If two matches are scheduled at the same time it is not possible to broadcast both of them. Therefore, the number of highly attended matches carried out in each period shall be kept below a certain threshold leading to a balanced distribution of attractive matches over all periods.

We introduce parameter $a_{i,j}$ for all $i, j \in T, i \neq j$, indicating whether a match (i, j, p) with $p \in P$ is attractive ($a_{i,j} = 1$) or not ($a_{i,j} = 0$). Obviously, in most cases $a_{i,j} = a_{j,i}$ holds. However, $a_{i,j} > a_{j,i}$ might make sense if i is a team having medium strength and j is a top level team. If j plays away it probably has an easy win but if the match is carried out at i's venue it might be a close and exciting match. Furthermore, we employ the parameter a_{max} as an upper bound for the number of attractive matches scheduled per period. The limitation is represented by $n-1$ constraints (18).

$$\sum_{i \in T} \sum_{j \in T \setminus \{i\}} a_{i,j} x_{i,j,p} \leq a_{max} \quad \forall p \in P \tag{18}$$

4 Fairness Constraints

In this section we examine constraints assuring some kind of fairness in RRTs. Such constraints are at the organizers' disposal (and hence can be adopted, modified, or dropped). Again, the Single RRT-IP serves as starting point.

4.1 Breaks

4.1.1 Minimum Number of Breaks

Team i is said to have a break in period p if i plays twice at home or away, respectively, in consecutive periods $p-1$ and p. We distinguish between home-breaks and away-breaks according to the venue of both matches. The most prominent goal considering breaks is minimizing the number of occurence. In de Werra [6] the number of breaks is shown to be no less than $n-2$. Note that $P^{\geq 2}$ denotes the set of periods excluding the first one. We represent the requirement of a minimum number of breaks by $2n(n-2)+1$ constraints employing $n(n-2)$ additional binary variables $br_{i,p}$ for all $(i,p) \in T \times P^{\geq 2}$.

$$\sum_{j \in T \setminus \{i\}} (x_{i,j,p-1} + x_{i,j,p}) - br_{i,p} \leq 1 \qquad \forall i \in T, p \in P^{\geq 2} \tag{19}$$

$$\sum_{j \in T \setminus \{i\}} (x_{j,i,p-1} + x_{j,i,p}) - br_{i,p} \leq 1 \qquad \forall i \in T, p \in P^{\geq 2} \tag{20}$$

$$\sum_{i \in T} \sum_{p \in P} br_{i,p} \leq n-2 \tag{21}$$

$$br_{i,p} \in \{0,1\} \quad \forall i \in T, p \in P^{\geq 2} \tag{22}$$

Constraints (19) and (20) force $br_{i,p}$ to be equal to 1 if team i has a home-break or an away-break in period p, respectively. Since we assure the minimum number of breaks by restriction (21) $br_{i,p}$ is equal to 1 if and only if i has a break in period p.

Additionally, we consider a common requirement in real world leagues: no break must occur in the second period. The reason for this is that breaks are considered as some kind of disturbance in the season's regularity. Disturbance shall be avoided at the season's beginning to guarantee a fair start. Breaks in the second period can be eliminated by letting $br_{i,2} = 0$ for each $i \in T$.

Note that a generation scheme is known for single RRTs having the minimum number of breaks, see de Werra [6]. However, single RRTs resulting from it are restricted to a specific class of RRTs (equivalent to a canonical oriented one-factorization) and, therefore, do not suffice if cost minimization is required.

4.1.2 One Break per Team

Although minimizing the number of breaks is most popular it is reasonable to think about arranging exactly one break for each team for fairness reasons as proposed in de Werra [9]. This can be covered using $2n(n-2)$ distinct variables $hbr_{i,p}$ and $abr_{i,p}$ representing the

decision about a home-break and an away-break, respectively, of team i in period $p \in P^{\geq 2}$ in $n(n-1)$ constraints.

$$\sum_{j \in T \setminus \{i\}} (x_{j,i,p-1} + x_{j,i,p}) + hbr_{i,p} - abr_{i,p} = 1 \qquad \forall i \in T, p \in P^{\geq 2} \tag{23}$$

$$\sum_{p \in P} (hbr_{i,p} + abr_{i,p}) = 1 \qquad \forall i \in T \tag{24}$$

$$hbr_{i,p} \in \{0,1\} \quad \forall i \in T, p \in P^{\geq 2} \tag{25}$$

$$abr_{i,p} \in \{0,1\} \quad \forall i \in T, p \in P^{\geq 2} \tag{26}$$

In order to illustrate constraint (23) we consider three cases. If team i has a home-break in period p the term in brackets equals 0 and, therefore, $hbr_{i,p}$ and $abr_{i,p}$ are forced to 1 and 0, respectively. If i has an away-break in period p the term in brackets sums up to 2 letting $hbr_{i,p} = 0$ and $abr_{i,p} = 1$. If i has no break in period p, finally, $hbr_{i,p} = abr_{i,p} = 0$ holds taking constraint (24) into account. Trivially, constraint (24) sets the number of breaks to 1 for each team $i \in T$.

4.2 Opponents' Strengths

Teams have different strengths, indicated for instance through the score obtained in the previous competition. Hence, two matches of team $i \in T$ might be distinctly exhaustive for i depending on the particular opponent. Assume that the teams are partitioned into a set of strength groups S. Each single team is contained in exactly one strength group $S' \in S$. Thus $S', S'' \subset T$, $|S' \cap S''| = 0$, and $\bigcup_{S' \in S} S' = T$ with $S', S'' \in S, S' \neq S''$, holds. For the sake of convenience we assume that $|T| \mod |S| = 0$ and that $|S'| = |S''|$ for all $S', S'' \in S$. Furthermore, each strength group contains teams having indices $k\frac{|T|}{|S|} + 1$ to $(k+1)\frac{|T|}{|S|}$ for $k \in \mathbb{N}_0, k < |S|$.

In order to avoid a burden of competition considered to be too high (and hence unfair) for some teams we claim that the matches of each team against teams of a specific strength group shall be distributed as even as possible over the tournament. In the following we present four different ways to enforce some kind of fairness considering the idea of strength groups.

An interesting question from a combinatorial point of view arises for each of these structures: for which numbers of strength groups $|S|$ a given number of teams $|T|$ is able to carry out a single RRT obeying one of the constraints concerning strength groups outlined in the following? This question is addressed in detail in Briskorn [3].

4.2.1 Changing Strength Groups

The most unpleasant case corresponds to two matches of team $i \in T$ in consecutive periods against teams belonging to the same strength group. See table 3 for an example considering team 1 in a league with 8 teams and 4 strength groups. Apparently, there is no change in the opponent's strength group between periods 3 and 4.

We employ binary variables $es^c_{i,p}$ to indicate the occurrence of such a case, that is, team $i \in T$ playing against teams of the same strength group in periods $p-1$ and p.

period	1	2	3	4	5	6	7
match	1-2	7-1	1-5	6-1	1-3	8-1	1-4
opponent's group	1	4	3	3	2	4	2

Table 3: Opponents' Strength Groups Not Changing

$$\sum_{j \in S', j \neq i} (x_{i,j,p-1} + x_{j,i,p-1} + x_{i,j,p} + x_{j,i,p}) \tag{27}$$

$$-es^c_{i,p} \leq 1 \qquad \forall i \in T, S' \in S, p \in P^{\geq 2}$$

$$\sum_{p \in P^{\geq 2}} es^c_{i,p} \leq es^c_{max} \quad \forall i \in T \tag{28}$$

$$es^c_{i,p} \in \{0,1\} \quad \forall i \in T, p \in P^{\geq 2} \tag{29}$$

Constraint (27) forces binary variable $es^c_{i,p}$ to be equal to 1 if team i has two opponents of the same strength group in periods $p-1$ and p. Inequality (28) assures that the number of violations of the changing strength group requirement is limited to a given parameter es^c_{max} representing the maximum number of allowed occurrences. Table 4 represents a feasible timetable for $i = 1$ if we consider $es^c_{max} = 0$. Table 3 is feasible if and only if $es^c_{max} > 0$ holds.

period	1	2	3	4	5	6	7
match	3-1	1-2	4-1	5-1	1-7	6-1	1-8
opponent's group	2	1	2	3	4	3	4

Table 4: Opponents' Strength Groups Changing

4.2.2 Balanced Strength Groups

Obviously, the timetable given in table 4 does not provide a perfect distribution of opponents' strength groups over the tournament. Only teams of groups 1 and 2 are opponents in periods 1 to 3, opponents of groups 3 and 4 are restricted to periods 4 to 7. In order to get a more balanced distribution we present a formulation claiming that the number of occurences of each group as opponent of i is restricted to 1 in a time window containing $|S|$ consecutive periods.

$$\sum_{p'=\max(p-|S|+1,1)}^{\min(p,|P|)} \sum_{j \in S', j \neq i} (x_{i,j,p'} + x_{j,i,p'}) \qquad \forall i \in T, S' \in S, \tag{30}$$

$$-es^b_{i,p} \leq 1 \qquad p \in P^{\geq 2} \cup_{p''=1}^{|S|-2} \{|P| + p''\}$$

$$\sum_{p \in P^{\geq |S|}} es^b_{i,p} \leq es^b_{max} \quad \forall i \in T \tag{31}$$

$$es^b_{i,p} \in \mathbb{N}_0 \qquad \forall i \in T, p \in P^{\geq 2} \tag{32}$$

Constraint (30) sets $es^b_{i,p}$ to the number by which the occurence of strength group S' is higher than 1 in periods $p - |S| + 1$ to p (the time window under consideration). Note that this affects $es^b_{i,p}$ for more periods the closer the particular groups' appearances are. For instance consider tables 5 and 6.

period	1	2	3	4	5	6	7
match	3-1	1-5	7-1	2-1	1-8	6-1	1-4
opponent's group	2	3	4	1	4	3	2

Table 5: Opponents' Strength Groups Not Balanced (Distance = 2)

period	1	2	3	4	5	6	7
match	3-1	1-5	7-1	2-1	1-4	8-1	1-6
opponent's group	2	3	4	1	2	4	3

Table 6: Opponents' Strength Groups Not Balanced (Distance = 3)

In table 5 opponent group 4 appears in periods 3 and 5. Both periods are taken into account in $es'_{i,5}$ as well as in $es'_{i,6}$. In table 6 the same group appears at periods 3 and 6 a fact which is taken into account only in $es^b_{i,6}$. Consequently, a closer appearance of the same group results in a higher value on the left hand side of (31) and, hence, is considered worse.

Note that periods 1 and 2 as well as periods $n-2$ and $n-1$ are covered only by one time window of length $|S|$. In order to recognize closer appearances in those periods, too, we establish $es^b_{i,p}$ for $2 \leq p < |S|$ and artificial periods $|P| < p \leq |P| + |S| - 2$. These time windows can not have length $|S|$ since they would reach beyond periods 1 and $|P|$, respectively. Instead, they are limited by the first and the last period, respectively, which is formally represented in the limits of the first sum of restriction (30). Doing so we obtain a perfectly balanced distribution of opponents' strength groups when we set $es^b_{max} = 0$ in (31), see table 7.

period	1	2	3	4	5	6	7
match	3-1	1-5	7-1	2-1	1-4	6-1	1-8
opponent's group	2	3	4	1	2	3	4

Table 7: Opponents' Strength Groups Balanced

4.2.3 Equally Unchanging Strength Groups

Taking into account other constraints the formulation introduced in section 4.2.1 might turn out to be too restrictive. Therefore, we introduce a formulation allowing a number of violations of constraints (27) to (29). Due to fairness reasons the number of violations shall be identical for each team. We introduce integer variables $es^c_{i,p,S'}$ for all $i \in T, p \in P^{\geq 2}, S' \in S$, counting the number of matches of i against teams of S' in periods $p-1$ and p. Consequently, the value of variables $es^c_{i,p,S'}$ is restricted to the set $\{0, 1, 2\}$, see (37). Furthermore we employ binary variables $f^c_{i,p,S'}$ indicating if there is a violation of changing strength groups attributed to team i and strength group S' in periods $p-1$ and p.

$$\sum_{j \in S', j \neq i} (x_{i,j,(p-1)} + x_{j,i,(p-1)}$$

$$+x_{i,j,p} + x_{j,i,p}) - es^c_{i,p,S'} = 0 \qquad \forall i \in T, p \in P^{\geq 2}, S' \in S \tag{33}$$

$$es^c_{i,p,S'} - f^c_{i,p,S'} \leq 1 \qquad \forall i \in T, p \in P^{\geq 2}, S' \in S \tag{34}$$

$$es^c_{i,p,S'} - 2f^c_{i,p,S'} \geq 0 \qquad \forall i \in T, p \in P^{\geq 2}, S' \in S \tag{35}$$

$$\sum_{p \in P^{\geq 2}} \sum_{S' \in S} f^c_{i,p,S'} = f^c_{fix} \qquad \forall i \in T \tag{36}$$

$$es^c_{i,p,S'} \in \mathbb{N}_0^{\leq 2} \qquad \forall i \in T, p \in P^{\geq 2}, S' \in S \tag{37}$$

$$f^c_{i,p,S'} \in \{0,1\} \qquad \forall i \in T, p \in P^{\geq 2}, S' \in S \tag{38}$$

Equation (33) initializes $es^c_{i,p,S'}$. Constraints (34) and (35) force $f^c_{i,p,S'}$ to be equal to 1 if and only if i has two matches against teams of S' in periods $p-1$ and p which is a violation of the changing strength group requirement. Constraint (36) sets the number of violations to the value of a given integer parameter f^c_{fix} for each team. Instead of this we can use f^c_{fix} as an integer variable. Then, the number of violations is not fixed a priori but is part of the solution of an instance.

4.2.4 Equally Unbalanced Strength Groups

In analogy to section 4.2.3 we allow violations according to the balanced strength group concept introduced in section 4.2.2 in the following. For the sake of convenience we introduce P' to cover all periods and artificial periods as outlined in section 4.2.2: $P' = P^{\geq 2} \cup_{p''=1}^{|S|-2} \{|P|+p''\}$. Since each strength group shall appear exactly once in each time window of length $|S|$ we can measure violations by counting the number of groups not contained in a specific time window. Again, we employ variables $es^b_{i,p,S'}$ to count matches between i and teams of S' in the time window corresponding to p and variables $f^b_{i,p,S'}$ to indicate violations. Violations occur as a strength group being multiply represented. Since a violation according to team i and period p directly correspond to a strength group not providing an opponent of i in the time window corresponding to p it is not possible to sum up the differences

$$d_{i,p,S'} = \sum_{p'=\max(p-|S|+1,0)}^{\min(p,|P|)} \sum_{j \in S': j \neq i} (x_{i,j,p'} + x_{j,i,p'}) - 1.$$

Note that differences $d_{i,p,S'}$ for all S' sum up to 0 for a team $i \in T$ and a period $p \in P$ since the number of matches for each time window is fixed to $|S|$. Therefore, we count the number $d_{i,p}$ of strength groups not appearing in a time window $p - |S| + 1$ to p as an opponent of i as follows:

$$d_{i,p} = \left| \left\{ S' \in S : \sum_{p'=\max(p-|S|+1,0)}^{\min(p,|P|)} \sum_{j \in S': j \neq i} (x_{i,j,p'} + x_{j,i,p'}) = 0 \right\} \right|$$

$$\sum_{p'=\max(p-|S|+1,1)}^{\min(p,|P|)} \sum_{j\in S',j\neq i} (x_{i,j,p'} + x_{j,i,p'}) - es^b_{i,p,S'} = 0 \qquad \forall i \in T, p \in P', S' \in S \tag{39}$$

$$es^b_{i,p,S'} + f'_{i,p,S'} \geq 1 \qquad \forall i \in T, p \in P', S' \in S \tag{40}$$

$$es^b_{i,p,S'} + |S| f'_{i,p,S'} \leq |S| \qquad \forall i \in T, p \in P', S' \in S \tag{41}$$

$$\sum_{p\in P^{\geq 2}} \sum_{S'\in S} f'_{i,p,S'} = \sum_{k=1}^{|S|-2} k + f^b_{fix} \qquad \forall i \in T \tag{42}$$

$$es^b_{i,p,S'} \in \mathbb{N}_0^{\leq |S|} \qquad \forall i \in T, p \in P', S' \in S \tag{43}$$

$$f^b_{i,p,S'} \in \{0,1\} \qquad \forall i \in T, p \in P', S' \in S \tag{44}$$

Equation (39) sets $es'_{i,p,S'}$ to the number of matches of team $i \in T$ against teams of strength group S' in the time window corresponding to period p. Constraints (40) and (41) force $f^b_{i,p,S'}$ to be equal to 1 if and only if $es^b_{i,p,S'} = 0$ holds. Note that $\sum_{S'\in S} f^b_{i,p,S'} = d_{i,p}$. Equation (42) assures that the number of violations are equal to f^b_{fix} for each team. We have to consider that, obviously, in time windows corresponding to artificial periods as well as periods 1 to $|S| - 1$ there are strength groups missing. The time window corresponding to $es^b_{i,2,S'}$, for example, can not contain more than two opponents and, therefore, strength groups. We take care of the number of this inherent lack of strength groups by adding $\sum_{k=1}^{|S|-2} k$ to the number of violations f^b_{fix}. As mentioned in section 4.2.3 we can substitute f^b_{fix} by an integer variable so the number of violations is not fixed but is forced to be identical for all teams.

4.3 Teams' Preferences

Often teams have preferences to play at home or away in specific periods. In real world sports leagues there are lots of reasons for that. There might be a major regional event so the team located in the region wants to play at home in order to attract more visitors. A prominent example is Bayern Munich preferring to play at home during the Octoberfest. The other way round teams might prefer to play away during major events in order to unload the infrastructure. Last but not least teams might prefer to play away due to construction works at the team's stadium causing lowered seating capacity.

In order to represent a team i's preference in period p we introduce the trivalent parameter $pr_{i,p}$. If team i wants to play at home in period p we set $pr_{i,p}$ to 1, if it wants to play away $pr_{i,p}$ is set to -1; if team i has no preference at all we have $pr_{i,p} = 0$. Obviously, we deserve a reasonable mechanism to handle teams having different numbers of preferences. Therefore, we let the number of possible preferences per team be unlimited but we construct sports league schedules obeying exactly the same number cp of preferences for each team. We employ the binary variable $np_{i,p}$ in order to indicate a neglected preference of team i in period p.

$$pr_{i,p} \sum_{j \in T, j \neq i} x_{i,j,p} - pr_{i,p} \sum_{j \in T, j \neq i} x_{j,i,p} + 2np_{i,p} = |pr_{i,p}| \quad \forall i \in T, p \in P \tag{45}$$

$$\sum_{p \in P} np_{i,p} - \max\left(\sum_{p \in P} |pr_{i,p}| - cp, 0\right) = 0 \quad \forall i \in T \tag{46}$$

$$np_{i,p} \in \{0,1\} \quad \forall i \in T, p \in P \tag{47}$$

Equation (45) sets $np_{i,p}$ to 1 if and only if team i's preference in period p is neglected. In order to line (45) out in detail we distinguish three cases. First, if team i has a preference to play at home in period p ($pr_{i,p} = 1$) the right hand side is equal to 1 and the first two terms of the left hand side sum up to 1 if team i plays at home in period p ($pr_{i,p} = -1$, otherwise). Consequently, $np_{i,p} = 1$ if and only if team i's preference in period p is neglected. Second, if team i has a preference to play away in period p ($pr_{i,p} = -1$) the right hand side is equal to 1 and the first two terms of the left hand side sum up to 1 if team i plays away in period p ($pr_{i,p} = -1$, otherwise). Again, $np_{i,p} = 1$ if and only if team i's preference in period p is neglected. Third, if team i has no preference in period p the right hand side equals 0 and the first two terms of the left hand side sum up to 0 no matter where i plays in period p. Hence, $np_{i,p} = 0$. Equation (46) forces the number of neglected preferences for each team i to be equal to the number of preferences given by i minus the number of preferences to be obeyed. Obviously, restriction (46) can be formulated as the number of neglected preferences for each team to be no more than the maximum expression in order to guarantee a certain number of obeyed preferences.

5 Computational Study

5.1 Generating Problem Instances

In order to establish benchmarks for algorithms concerning various sports league scheduling problems we propose an advanced instance generator covering all issues introduced in sections 3 and 4. So far, the generator does not include further aspects such as, for example, travel distances. In the following we line out the specification of a problem instance in detail. Furthermore, we define parameters enabling us to influence the problem instances' characteristics.

Problem Size: The size of an instance is determined by the number n of teams. Recall that n must be even w.l.o.g. as shown in section 1.

Forbidden Matches: If the set of possible matches is to be restricted according to section 3.1 we need a parameter $\pi_{i,j,p}$ for each match indicating whether it is forbidden or not. We control the fraction of matches which are allowed by parameter P_π. P_π gives the probability for each match (i,j,p) to be possible ($\pi_{i,j,p} = 1$). When constructing an instance each $\pi_{i,j,p}$ is randomly chosen from $\{0,1\}$ according to P_π.

Stadium Availability: As described in section 3.1 stadium unavailability can be seen as a set of forbidden matches. However, we decided to employ a distinct parameter for guiding the number of times a stadium is not available which is expected to have an impact on run times. Parameter P_π^S gives the probability for a team i's stadium to be available in a specific period

p. The availability of of the stadium of each team in each period is chosen randomly from $\{0,1\}$ according to P_π^S.

Regions: If regions' capacities are considered we have to specify the set of regions R according to section 3.2. Each region is defined by both, the teams it contains and the maximum number of matches being allowed in this region per period, explicitly.

Highly Attended Matches: In order to study different constellations of attractive matches we employ two parameters in order to construct problem instances. First, P_a gives the probability for a match (i,j,p) to be attractive. Second, the maximum number of attractive matches per period a_{max} must be given explicitly. Reasonably, values for a_{max} are from $\{1,\ldots,\frac{n}{2}-1\}$. Parameter a_{max} being equal to 0 does not allow any attractive match at all while a_{max} being no less than $\frac{n}{2}$ does not define any restriction at all since there are exactly $\frac{n}{2}$ matches per period. When constructing the problem instance $a_{i,j}$ is randomly chosen from $\{0,1\}$ according to P_a.

Breaks: If breaks are to be considered we distinguish four classes of instances. First, the overall goal is either to minimize the number of breaks or to arrange exactly one break per team. Second, we allow and forbid, respectively, breaks to occur in the second period. Beside this choice there are no further parameters necessary.

Opponents' Strengths: In order to consider opponents' strengths three parameters must be defined: the mode of consideration, the set of strength groups, and the number of violations. The mode of consideration can be chosen from "changing", "balanced", "equally unchanging", and "equally unbalanced". Following the concept outlined in section 4.2 the set of strength groups can be defined by specifying the number of strength groups. Additionally, the number of violations according to the changing strength groups and balanced strength groups has to be given explicitly.

5.2 Computational Results

This study was carried out employing Cplex 9.0 on a 3.8 GHz Pentium 4 personal computer using 3 GBs of RAM. Cplex was executed with default parameter settings. We state average computation times for runs only which terminated, number of instances tested and the number of instances a solution could be found for versus the number of problems having no solution at all. Note that runs not leading to an optimal solution or to the recognition of infeasibility have been aborted because of lack of memory. In what follows "i.", "s.f.", "n.s.", and "r.t." denote "number of instances", "number of instances solutions are found for", "number of solutions proven to be infeasible", and "average run times" in seconds (for the instances solved to optimality), respectively.

5.2.1 Basic Problem

First, we observe the run time behavior when solving the basic single RRT problem without any additional restrictions. Results are provided in table 8.

We clearly observe an explosion of run times as problem sizes grow. While run times for problems with up to 12 teams are rarely worth mentioning they extremely raise for larger instances. Instances with 16 teams can be solved to optimality in 47 minutes on average but

n	i.	s.f.	r.t.
6	20	20	0.01
8	20	20	0.05
10	20	20	0.37
12	20	20	2.03
14	20	20	34.43
16	20	20	2811.18
18	3	1	74601.00

Table 8: Computational results for basic single RRT problem

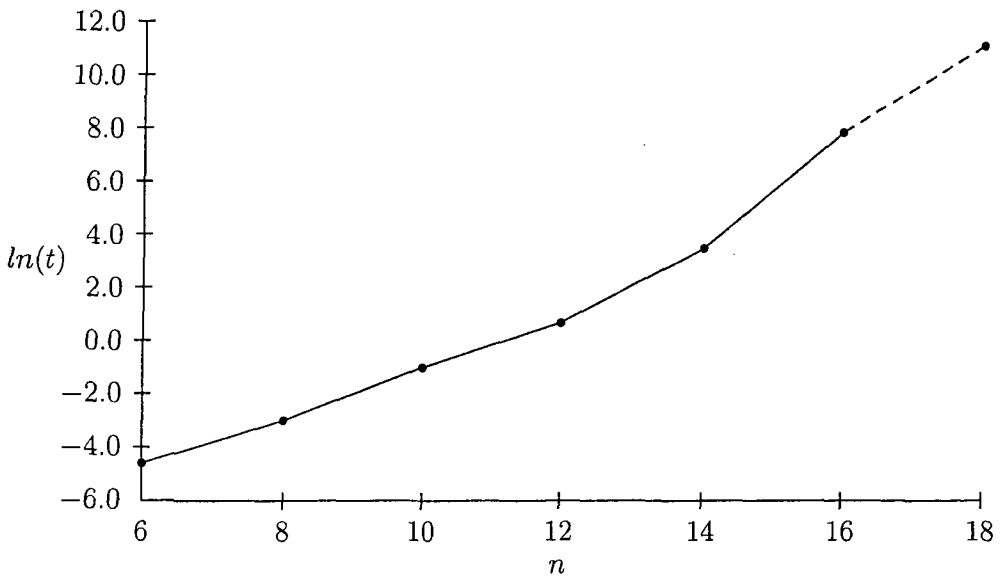

Figure 1: Run time behavior for basic single RRT problem

only one instance with 18 teams could be solved to optimality. This took more than 20.5 hours. Cplex ran out of memory while solving the others.

Figure 1 illustrates the exponential run time behavior. Note that the logarithm $\ln(t)$ of run time t is shown. The function progresses almost linearly. The last section is drawn as dashed line since there was only one instance indicating this run time. Nevertheless, this time fits into the linear scheme recognized for smaller instances. Those runs suffering from lack of memory were aborted after about 14 hours.

5.2.2 Forbidden Matches

As described in section 3.1 forbidden matches cut down solution space since RRTs must not contain a forbidden match. This might lead to decreasing run times. On the other hand finding feasible solutions gets more difficult and cost oriented branching mechanisms might get stuck in infeasible paths more likely. Hence, we have two counterrotating effects influencing run times in comparison to those presented in section 5.2.1. In order to study the impact of those effects we tested instances with 6 to 18 teams and probabilities for matches to be allowed

of $P_\pi \in \{0.9, 0.8, 0.7, 0.6, 0.5\}$. Results are lined out in table 9. For each class of instances we give size, number of instances, number of instances solved to optimality, and number of infeasible instances, respectively.

	P_π =0.9		P_π =0.8		P_π =0.7	
n	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.
6	20/20/0	0.01	20/20/0	0.01	20/20/0	0.01
8	20/20/0	0.05	20/20/0	0.03	20/20/0	0.03
10	20/20/0	0.39	20/20/0	0.35	20/20/0	0.32
12	20/20/0	1.84	20/20/0	2.30	20/20/0	1.80
14	20/20/0	44.16	20/20/0	32.95	20/20/0	34.53
16	20/20/0	2236.30	20/20/0	2252.25	20/20/0	2647.91
18	3/0/0	———	3/1/0	18166.90	3/2/0	43820.15

	P_π =0.6		P_π =0.5	
n	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.
6	20/19/1	0.01	20/17/3	0.01
8	20/20/0	0.04	20/20/0	0.04
10	20/20/0	0.42	20/20/0	0.40
12	20/20/0	2.10	20/20/0	2.03
14	20/20/0	41.41	20/20/0	30.15
16	20/20/0	3898.66	20/20/0	3094.34
18	3/1/0	27288.90	3/1/0	29698.80

Table 9: Computational results for forbidden matches

Up to 10 teams we observe almost equal run times for all P_π not different from those given for the basic problem. For larger instances run times do not provide a clear idea about the influence of the number of teams and probability P_π on run times. For example while $P_\pi = 0.9$ leads to largest run times for instances having 14 teams it leads to the smallest run times for instances having 16 teams. Unfortunately, we can not observe any systematic correlation, here. However, characteristics of average run times' behavior for instances with forbidden matches do not severely differ from those presented for the basic problem shown in figure 1, therefore, we restrain to give them here explicitly.

We emphasize that infeasibility of instances is rather odd even if 50% of matches are forbidden. Infeasibility was detected for only a few of the smallest of instances.

5.2.3 Stadium Availability

As outlined in section 3.1 we represent stadiums unavailability as forbidden matches. In addition to the results in section 5.2.2 we have a further look on run times when stadium availability is considered because of the high relevance in real world RRTs.

We create test instances having 6 to 18 teams and $P_\pi^S \in \{0.9, 0.8, 0.7\}$. Results are given in table 10.

Trivially, probability of instances' infeasibility is higher if probability of stadium availability is lower. This thought is confirmed by the fraction of infeasible instances of each problem class.

n	P_π^S =0.9		P_π^S =0.8		P_π^S =0.7	
	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.
6	10/10/0	0.01	10/10/0	0.01	10/8/2	0.01
8	10/10/0	0.03	10/10/0	0.03	10/6/4	0.03
10	10/10/0	0.43	10/9/1	0.36	10/5/5	0.17
12	10/10/0	2.33	10/10/0	2.02	10/7/3	1.42
14	10/10/0	31.31	10/10/0	27.48	10/9/1	9.71
16	10/10/0	2100.76	10/9/1	1308.86	10/5/5	1085.94
18	3/1/0	64987.30	3/2/0	29126.90	3/3/0	30339.27

Table 10: Computational results for stadium availability

No instance having $P_\pi^S = 0.9$ is infeasible while for each problem size there are infeasible instances having $P_\pi^S = 0.7$. Comparing tables 9 and 10 we conclude that probability of instance's infeasibility is clearly higher when considering stadium availability instead of arbitrarily forbidden matches even if $P_\pi^S = P_\pi$.

By contrast decreasing probability of stadium availability lowers run time requirements. This can be observed for $n < 18$. Note that we can think of the basic problem as a problem having $P_\pi^S = 1$. Then, results from table 8 fit into this observation since run times according to the basic problem are larger than those in table 10 for $n > 12$. For $n = 18$ run times may not be representative since we could only test very few instances. However, three out of three instances with $P_\pi^S = 0.7$ and $n = 18$ terminated. This fraction decreases with increasing P_π^S which may be an indicator for larger run time requirements.

Furthermore, considering stadium availability seems to lead to lower run times than general forbidden matches do. Note that instances with P_π^S can be considered as special cases of instances with $P_\pi^S = P_\pi^S$. However, run times for instances considering stadium availability are significantly lower than for general forbidden matches for $12 < n < 18$. Run times for smaller instances are almost equal.

5.2.4 Highly Attended Matches

When considering limitation of the number of highly attended matches per period we expect to experience the same two effects as lined out before: reduction of solution space versus difficulty of finding feasible solutions.

We created instances where each match has probability $P_a = 0.2$ to be attractive. We vary a_{max} depending on the instances' sizes. Note that for $n = 10$ and $P_a = 0.2$ there can not be less than one attractive match per period on average. Therefore, we increase a_{max} from 1 to 2.

Obviously, a_{max} is differently restrictive for instances having 6 and 8 teams and having 10, 12, 14, 16, and 18 teams, respectively. This leads to a distortion of run times. However, run times in table 11 are comparable to those of the basic problem in table 8. We clearly observe larger run times when considering highly attended matches.

n	P_a	a_{max}	i./s.f./n.s.	r.t.
6	0.2	1	10/10/0	0.01
8	0.2	1	10/10/0	0.09
10	0.2	2	10/10/0	0.29
12	0.2	2	10/10/0	2.78
14	0.2	2	10/10/0	96.90
16	0.2	2	10/10/0	16925.70
18	0.2	2	3/0/0	—

Table 11: Computational results for highly attended matches

5.2.5 Regions' Capacity

As outlined in section 3.2 we can think of a huge amount of constellations taking regions into account. For a given size of instances n there can be various numbers of regions, various sizes of regions, and various assignments of teams to regions (which can influence run times due to the different costs).

We exemplarily create three classes R_2, R_3, and R_4 of instances. In each class we have a given number of regions of similar size. The number of matches in a specific region is limited by $C_{R'} = \left\lfloor \frac{|R'|}{2} \right\rfloor + 1$. Since $C_{R'} \geq \left\lceil \frac{|R'|}{2} \right\rceil$ must hold (see section 3.2) the limitation in R_2, R_3, and R_4 is quite restrictive.

In R_2 there are two disjoint regions of size $\frac{n}{2} - 1$ and $\frac{n}{2} + 1$, respectively. In R_3 we have three disjoint regions having minimum size of $\lfloor \frac{n}{3} \rfloor$. $n \mod 3$ teams have size $\lfloor \frac{n}{3} \rfloor + 1$. In R_4 there are four disjoint regions having minimum size of $\lfloor \frac{n}{4} \rfloor$. $n \mod 4$ teams have size $\lfloor \frac{n}{4} \rfloor + 1$.

n	$\lvert R' \rvert$	$C_{R'}$	i./s.f.	r.t.	$\lvert R' \rvert$	$C_{R'}$	i./s.f.	r.t.
6	2/4	2/3	10/10	0.01	2/2/2	2/2/2	10/10	0.01
8	3/5	2/3	10/10	0.09	2/3/3	2/2/2	10/10	0.08
10	4/6	3/4	10/10	0.40	3/3/4	2/2/3	10/10	0.83
12	5/7	3/4	10/10	3.65	4/4/4	3/3/3	10/10	4.82
14	6/8	4/5	10/10	81.35	4/5/5	3/3/3	10/10	345.28
16	7/9	4/5	10/10	4080.78	5/5/6	3/3/4	10/5	32663.66
18	8/10	5/6	3/0	—	6/6/6	4/4/4	3/0	—

n	$\lvert R' \rvert$	$C_{R'}$	i./s.f.	r.t.
6	1/1/2/2	1/1/2/2	10/10	0.02
8	2/2/2/2	2/2/2/2	10/10	0.03
10	2/2/3/3	2/2/2/2	10/10	0.83
12	3/3/3/3	2/2/2/2	10/10	16.21
14	3/3/4/4	2/2/3/3	10/10	546.06
16	4/4/4/4	3/3/3/3	10/10	20775.20
18	4/4/5/5	3/3/3/3	3/0	—

Table 12: Computational results for regions

Each single class of problem instances proofed to be feasible. Almost all instances could be solved to optimality. Each instance having 18 teams could not be solved to optimality. However, for those instances not solved to optimality feasible solutions were found.

With respect to the run times we can conclude that mostly a larger number of regions means higher run times. We observe exceptions from this rule for instances having 16 teams and 3 and 4 regions, respectively. A possible explanation for this effect is that limitation for the number of matches in a specific region is much more restrictive in those instances having 3 regions. For example for instances having 16 teams and 3 regions the number of matches per region is restricted to about 0.62 times the number of teams in a region on average. This rate amounts to 0.75 for instances having 16 teams and 4 regions and is, therefore, less restrictive. Furthermore, this effect elucidates why only half of the instances having 16 teams and 3 regions could be solved to optimality.

In comparison with the basic problem we can conclude that run times clearly rise when regions capacities are considered.

5.2.6 Breaks

We study problems incorporating requirements considering breaks as introduced in section 4.1. In detail we solve problems requiring a minimum number of breaks (allowing and forbidding breaks in the second period) and having exactly one break per team. Run times are lined out in table 13.

n	min no	min no, not 2nd	one b. per team
6	0.23	0.21	0.25
8	26.79	26.63	38.44
10	—	—	—

Table 13: Computational results for breaks

Clearly, allowing and forbidding breaks in the second period has no impact on run times when a minimum number of breaks is required. For $n = 10$ Cplex aborted due to lack of memory after about 15 hours of run time (in most cases without even a single feasible solution). This, first of all, shows the enormous increase of run times for solving these problems to optimality. Second, it gives an idea of the difficulty to find feasible solutions when cost oriented branching is employed instead of the standard generation scheme mentioned in section 4.1.

Run times for instances requiring exactly one break per team are even higher. Again, we can not solve instances with more than $n = 8$ teams. Since problems having $n = 6$ teams and forbidding breaks in the second period have no feasible solution at all we refuse to line out results according to the missing instance class.

5.2.7 Opponents' Strengths

In section 4.2 four variants to consider a team's opponents' strengths in order to establish fairness are proposed. We tested instances having up to $n = 18$ teams for all variants. The number of strength groups is set to 2 and $\frac{|T|}{2}$, respectively.

For "changing" and "balanced" instances are fully specified by parameters given above. For "equally unchanging" and "equally unbalanced" we additionally set the number of violations to 1, which means each team must violate the opponent strength constraint exactly once.

	$\lvert S\rvert = 2$		$\lvert S\rvert = \frac{n}{2}$	
n	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.
6	10/0/10	—	10/0/10	—
8	10/10/0	0.02	10/10/0	0.34
10	10/0/10	—	10/10/0	4.59
12	10/10/0	0.18	10/10/0	233.64
14	10/0/10	—	10/10/0	11559.40
16	10/10/0	10.33	3/0/0	—
18	10/0/10	—	—	—

Table 14: Computational results for changing opponents' strength

In tables 14 to 17 results for all classes of instances are provided. Referring to the question posed in section 4.2 we can identify problem classes considering changing opponents' strength being infeasible due to values of n and $|S|$ according to table 14. For $n \leq 18$, $(n-2) \mod 4 = 0$, and $|S| = 2$ we observe infeasibility. We conjecture this to be true for $n > 18$. Furthermore, $n = 6$ and $|S| = 3$ leads to infeasibility as well

Run times for feasible instances having $|S| = 2$ are significantly lower than for basic ones of corresponding size n. To the contrary run times are significantly higher for $|S| = \frac{n}{2}$ in comparison to both, the basic instance and instance with "changing" opponent strength and $|S| = 2$. Consequently, instances having 16 teams and more can not be solved to optimality.

	$\lvert S\rvert = 2$		$\lvert S\rvert = \frac{n}{2}$	
n	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.
6	10/0/10	—	10/0/10	—
8	10/10/0	0.03	10/10/0	0.03
10	10/0/10	—	10/0/10	—
12	10/10/0	0.61	10/10/0	0.50
14	10/0/10	—	10/0/10	—
16	10/10/0	10.09	10/0/10	9.26
18	10/0/10	—	10/0/10	—

Table 15: Computational results for balanced opponents' strength

Inspecting table 15 again we identify problem classes being infeasible. Note that balanced opponents' strength structure is a special case of changing opponents' strength structure. Therefore, it is straightforward that n and $|S|$ is infeasible according to balanced opponent's strength if n and $|S|$ is infeasible according to changing opponent's strength. However, we find some instance classes being feasible according to changing opponent's strength but being infeasible according to balanced opponent's strength: For $n \leq 18$, $(n-2) \mod 4 = 0$, and $|S| = \frac{n}{2}$ we observe infeasibility. Again, we conjecture this be valid for $n > 18$.

Run times according to both, $|S| = 2$ and $|S| = \frac{n}{2}$, are significantly lower than for the basic

problem. Run times for $|S| = \frac{n}{2}$ are slightly lower than those for $|S| = 2$ which is in contrast to the relation observed for changing opponents' strength.

n	$\lvert S\rvert = 2$		$\lvert S\rvert = \frac{n}{2}$	
	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.
6	10/0/10	—	10/0/10	—
8	10/10/0	0.76	10/10/0	8.26
10	10/0/10	—	10/10/0	3116.31
12	10/10/0	195.01	3/0/0	—
14	3/0/0	—	—	—

Table 16: Computational results for equally unchanging opponents' strength

Run times according to equally unchanging opponents' strength are given in table 16. Here we observe remarkably higher run times compared with changing opponents' strength given in table 14. The reason for this might be the larger number of binary variables necessary to represent the equally unchanging opponents' strength. Furthermore, integer variables not restricted to binary values are incorporated. As far as table 16 provides insights into this topic exactly the same classes of problem instances seem to be infeasible as are for changing opponents' strength.

Finally, run times according to equally unbalanced opponents' strength are given in table 17. Run times are clearly higher than for balanced opponents' strength and equally unchanging

n	$\lvert S\rvert = 2$		$\lvert S\rvert = \frac{n}{2}$	
	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.
6	10/0/10	—	10/0/10	—
8	10/10/0	0.63	10/10/0	2.40
10	10/0/10	—	10/0/10	—
12	10/10/0	639.20	10/10/0	17290.30
14	3/0/0	—	3/0/0	—

Table 17: Computational results for equally unbalanced opponents' strength

opponents' strength given in tables 15 and 16, respectively. Again, results give only an slight idea of classes of instances being infeasible. However, as it was the case for changing and equally unchanging opponents' strength there seems to be no difference according to problems feasibility between balanced and equally unbalanced opponents' strength.

5.2.8 Teams' Preferences

We consider six classes of instances according to section 4.3: teams specify 1-2, 1-3, and 2-4 preferences, and exactly or at least 1, 1, and 2 preferences, respectively, have to be considered. Obviously, one can think of more preferences if the number of teams (and, therefore, the number of periods) is larger. For the sake of comparability we refuse to do so. Results are given in table 18.

n	1-2, exactly 1		1-3, exactly 1		2-4, exactly 2	
	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.
6	10/10/0	0.01	10/10/0	0.02	10/10/0	0.02
8	10/10/0	0.09	10/10/0	0.14	10/10/0	0.20
10	10/10/0	0.50	10/10/0	0.86	10/10/0	1.20
12	10/10/0	3.14	10/10/0	4.01	10/10/0	5.84
14	10/10/0	66.05	10/10/0	157.36	10/10/0	359.98
16	10/10/0	6192.64	10/10/0	12223.80	10/10/0	21002.60
18	3/1/0	156683.00	3/0/0	—	3/0/0	—

n	1-2, at least 1		1-3, at least 1		2-4, at least 2	
	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.	i./s.f./n.s.	r.t.
6	10/10/0	0.01	10/10/0	0.02	10/10/0	0.01
8	10/10/0	0.05	10/10/0	0.06	10/10/0	0.09
10	10/10/0	0.39	10/10/0	0.48	10/10/0	0.66
12	10/10/0	2.81	10/10/0	3.05	10/10/0	2.90
14	10/10/0	90.45	10/10/0	67.22	10/10/0	53.91
16	10/10/0	3596.75	10/10/0	2841.99	10/10/0	13205.00
18	3/0/0	—	3/0/0	—	3/0/0	—

Table 18: Computational results for teams' preferences

For each single problem class and size we can conclude that run times are significantly higher than for the basic problem having identical size. Furthermore, run times for each problem class considering a given exact number of preferences to be granted are higher than for the corresponding class considering a given minimum number of preferences to be granted. This effect might result from the difficulty to find feasible solutions. Obviously, the set of solutions having an exact number p' of granted preferences per team is a subset of the set of solutions having p' as a minimum number of granted preferences.

As we can see increasing the number of preferences and increasing the number of preferences to be granted leads to higher run times if an exact number of preferences to be fulfilled is given. Again, the reason for this probably is the difficulty to find feasible solutions: increasing number of preferences leads to a rising number of preferences which must be neglected. This effect does not come into play if we consider a minimum number of preferences to be fulfilled. Instead, by increasing the number of preferences there is more freedom to choose the number of preferences to be fulfilled. Therefore, there is a tendency that run times are lower if 1 to 3 preferences are given in comparison with 1 to 2 preferences.

6 Summary and Outlook

In this paper we pick up several prominent real world requirements in the context of RRT scheduling. Furthermore, we substantiate requirements related to fairness which mostly have been proposed in abstract terms in literature so far. We formally define the requirements by the means of IP modelling techniques.

Moreover, we studied the run time behavior resulting of the optimization models using Cplex. We observe exponential run time behavior for nearly all variations of the minimum cost single RRT problem. Therefore, run times are exorbitant as soon as problem sizes grow relevant for real world problems except. We detect two exceptions from this rule: If we consider changing opponents' strength groups (as far as $|S|$ is small) and balanced opponents' strength groups, respectively, run time remains manageable. Hence, these variants might serve as basis for real world problems (mostly $n \leq 20$) although the corresponding optimization problems are NP-hard as shown in [3].

In this paper we have studied many variants of RRTs having relevance for real world leagues. Nevertheless, there are further variants and modelling techniques to be considered. For example, the type of variables employed in this paper is rather straightforward. There are alternatives which might lead to better LP-relaxations and, therefore, to a smaller amount of CPU time required.

References

[1] T. Bartsch. *Sportligaplanung – Ein Decision Support System zur Spielplanerstellung (in German)*. Deutscher Universitätsverlag, Wiesbaden, 2001.

[2] T. Bartsch, A. Drexl, and S. Kröger. Scheduling the Professional Soccer Leagues of Austria and Germany. *Computers & Operations Research*, 33:1907–1937, 2006.

[3] D. Briskorn. Combinatorial Properties of Strength Groups in Round Robin Tournaments. *Working Paper*, 2006.

[4] D. Briskorn, A. Drexl, and F. C. R. Spieksma. Round Robin Tournaments and Three Index Assignment. *Working Paper*, 2006.

[5] P. Brucker and S. Knust. *Complex Scheduling*. Springer, Berlin, 2006.

[6] D. de Werra. Geography, Games and Graphs. *Discrete Applied Mathematics*, 2:327–337, 1980.

[7] D. de Werra. Scheduling in Sports. In P. Hansen, editor, *Studies on Graphs and Discrete Programming*, pages 381–395. North-Holland, Amsterdam, The Netherlands, 1981.

[8] D. de Werra. Minimizing Irregularities in Sports Schedules using Graph Theory. *Discrete Applied Mathematics*, 4:217–226, 1982.

[9] D. de Werra. On the Multiplication of Divisions: the Use of Graphs for Sports Scheduling. *Networks*, 15:125–136, 1985.

[10] D. de Werra. Some Models of Graphs for Scheduling Sports Competitions. *Discrete Applied Mathematics*, 21:47–65, 1985.

[11] D. de Werra, T. Ekim, and C. Raess. Construction of Sports Schedules with Multiple Venues. *Discrete Applied Mathematics*, 154:47–58, 1985.

[12] A. Drexl and S. Knust. Sports League Scheduling: Graph- and Resource–Based Models. *Omega*, 35:465–471, 2007.

[13] J. A. M. Schreuder. Constructing Timetables for Sport Competitions. *Mathematical Programming Study*, 13:58–67, 1980.

[14] J. A. M. Schreuder. Combinatorial Aspects of Construction of Competition Dutch Professional Football Leagues. *Discrete Applied Mathematics*, 35:301–312, 1992.

[15] M. A. Trick. Integer and Constraint Programming Approaches for Round Robin Tournament Scheduling. In E. Burke and P. de Causmaecker, editors, *Proceedings of the 4th International Conference on the Practice and Theory of Automated Timetabling*, Lecture Notes in Computer Science 2740, pages 63–77. Springer, Berlin, Germany, 2003.