Fair Fixture: Minimizing Carry-Over Effects in Football Leagues

Dilek Günneç

Department of Industrial Engineering
Özyeğin University

with Ezgi Demir, BSH

July, 2019
MathSports, Athens

Outline

Motivation

- Sports is a multi-billion dollar industry.
- A key to attracting audience, broadcasting companies are interested in finding the best schedule with the highest advertisement revenues.
- Scheduling mainly consists of determining **opponents**, **date** and **venue** for each game.
- Objectives: Minimizing travelling distance, maximizing fairness, maximizing audience, etc.
- Finding only a feasible schedule with some basic requirements is a difficult problem.

Carry-over Effects

- Assume that i plays with j and j is strong. And next week when i plays with k, k will be advantageous as i will be tired from the previous week.
- Team k receives carry-over effects from team j.
- Carry-over matrix is represented with C, where c_{ij} is the number of times team i gives carry-over effects to team j and *carry-over effects value* of a league is determined by $\sum_{i \in T} \sum_{j \in T} c_{ij}^2$, where T is the set of teams.
- The lowest carry-over effects value possible in a single-round robin tournament[1] with n teams would be $n(n-1)$ if each team only followed every other team once.

[1] By definition, carry-over effects from the last week to the first is also counted.

Game Locations

- In double round robin tournaments, the first game is played at one of the team's "home (H)" town (the opponent is said to have the game "away (A)"), and the second game is played at the other team's home town.
- If there are two consecutive home or away games for a team, then it is called a "break" for that team. It can be shown that a league of n teams require a total of at least $n-2$ breaks in a half-season (D. De Werra, 1981); $(n-2)$ teams have 1 break and 2 teams have no breaks.
- Many of the European premier leagues (Belgium, England, Germany, Italy, Portugal, Spain, Turkey) follow this scheme where no team plays more than two home (or away) matches in a row (Goossens et. al. 2012).

Research Problem

- Creating a fair schedule
  - Minimize carry over effects.
  - While at most one break for each team in a half season
  - with no breaks in the first and last weeks.[2]

[2] In the aforementioned premier leagues it is customary to avoid any breaks during the first and last weeks.

Related Literature

- Minimizing carry-over effects (Russell, 1980)
  - Many approaches have been proposed starters-based [1], constraint programming [16, 10], a permutation scheme (an extension of the polygon method) [13], tabu search [11], iterated local search with multistart strategy [9] for the weighted carry-over effects problem and multi-phase scheduling [14] for the extended carry-over effects problem.
  - Lambrechts (2016) showed that the canonical form results in maximizing the carry-over effects.
- Minimizing the total number of breaks.
  - Well known constructive algorithm for minimizing number of breaks is the canonical form [2, 3, 4].
  - Separate the problem into three phases:
    1) Feasible pattern generation.
    2) Pattern matching with minimum number of breaks
    3) Team-Pattern assignment

Mathematical Model

- 3 sets of decision variables.
  - Patterns
    - x_{ip} is equal to 1 if team $i \in T$ follows pattern $p \in P$, 0 otherwise
    - y_p equals 1 if pattern $p \in P$ is used, 0 otherwise.
  - Opponent, week and location assignments:
    - x_{ijk} equals 1 if teams $i \in T$ and $j \in T$ play against each other in week $k \in K$ and 0 otherwise.
    - h_{ik} shows the location of the game and is equal to 1 if team $i \in T$ plays at home in week $k \in K$, and 0 if i plays away.
  - Carry-over effects
    - c_{ijk} equals to 1 if team $i \in T$ gives carry-over effect to team $j \in T$ in week $k \in K$, 0 otherwise
    - c_{ij} represents the total carry-over effect from team $i \in T$ to team $j \in T$
    - s_{pk} equals 1 if corresponding location for pattern p at week k is H, and 0 otherwise.

Mathematical Model

- Set of possible patterns can be determined beforehand.

Table: Pattern set for a league of 8 teams.

		Week						
		1	**2**	**3**	**4**	**5**	**6**	**7**
Pattern	1	A	H	A	H	A	H	A
	2	A	**H**	**H**	A	H	A	H
	3	A	H	**A**	**A**	H	A	H
	4	A	H	A	**H**	**H**	A	H
	5	A	H	A	H	**A**	**A**	H
	6	H	A	H	A	H	A	H
	7	H	**A**	**A**	H	A	H	A
	8	H	A	**H**	**H**	A	H	A
	9	H	A	H	**A**	**A**	H	A
	10	H	A	H	A	**H**	**H**	A

Mathematical Model

$$\text{INT:} \quad \text{Min} \quad \sum_{i=1}^{n}\sum_{j=1}^{n} c_{ij}^2 \tag{1}$$

subject to

$$\sum_{p=1}^{2n-6} x_{ip} = 1 \qquad i = 1, ..., n, \tag{2}$$

$$\sum_{i=1}^{n} x_{ip} = y_p \qquad p = 1, ..., 2n-6, \tag{3}$$

$$y_1 = 1, \tag{4}$$

$$y_p = y_{p+(n-3)} \qquad p = 1, ..., n-3, \tag{5}$$

$$x_{ijk} = x_{jik} \qquad i, j = 1, ..., n, \quad k = 1, ..., n-1, \tag{6}$$

$$x_{iik} = 0 \qquad i = 1, ..., n, \quad k = 1, ..., n-1, \tag{7}$$

Mathematical Model

$$\sum_{k=1}^{n-1} x_{ijk} = 1 \qquad i,j = 1,...,n, \quad i \neq j, \tag{8}$$

$$\sum_{j=1}^{n} x_{ijk} = 1 \qquad i = 1,...,n, \quad k = 1,...,n-1 \tag{9}$$

$$h_{ik} = \sum_{p=1}^{2n-6} x_{ip} s_{pk} \qquad i = 1,...,n, \quad k = 1,...,n-1, \tag{10}$$

$$x_{ijk} \leq 2 - (h_{ik} + h_{jk}) \qquad i,j = 1,...,n, \quad i \neq j \quad k = 1,...,n-1, \tag{11}$$

$$x_{ijk} \leq (h_{ik} + h_{jk}) \qquad i,j = 1,...,n, \quad i \neq j \quad k = 1,...,n-1, \tag{12}$$

Mathematical Model

$$x_{ilk} + x_{jl(k+1)} - 1 \leq c_{ijk} \qquad i,j,l = 1,...,n, \quad k = 1,...,n-2, \tag{13}$$

$$x_{il(n-1)} + x_{jl1} - 1 \leq c_{ij(n-1)} \qquad i,j,l = 1,...,n, \tag{14}$$

$$\sum_{k=1}^{n-1} c_{ijk} = c_{ij} \qquad i,j = 1,...,n, i \neq j \tag{15}$$

where $x_{ip}, y_p, x_{ijk}, h_{ik}, c_{ijk} \in \{0,1\}$
and $c_{ij} \geq 0 \quad i,j \in T \; k \in K, \; p \in P.$

Solution Method:Pattern and Opponent Assignment

Pattern Assignment

- Possible set of patterns can be identified at the outset.
- Find a feasible solution of the following model (PAT):

$$\sum_{p=1}^{2n-6} x_{ip} = 1 \qquad i = 1, ..., n,$$

$$\sum_{i=1}^{n} x_{ip} = y_p \qquad p = 1, ..., 2n-6,$$

$$y_1 = 1$$

$$y_p = y_{p+(n-3)} \qquad p = 1, ..., n-3,$$

$$h_{ik} = \sum_{p=1}^{2n-6} x_{ip} s_{pk} \qquad i = 1, ..., n, \quad k = 1, ..., n-1,$$

Solution Method: Pattern and Opponent Assignment

Opponent Assignment

- Find a feasible solution of the following model (OPP):

$$
\begin{aligned}
& x_{ijk} = x_{jik} && i, j = 1, ..., n, \quad k = 1, ..., n-1, \\
& x_{iik} = 0 && i = 1, ..., n, \quad k = 1, ..., n-1, \\
& \sum_{k=1}^{n-1} x_{ijk} = 1 && i, j = 1, ..., n, \quad i \neq j, \\
& \sum_{j=1}^{n} x_{ijk} = 1 && i = 1, ..., n, \quad k = 1, ..., n-1, \\
& x_{ijk} \leq 2 - (h_{ik} + h_{jk}) && i, j = 1, ..., n, \quad i \neq j \quad k = 1, ..., n-1, \\
& x_{ijk} \leq (h_{ik} + h_{jk}) && i, j = 1, ..., n, \quad i \neq j \quad k = 1, ..., n-1,
\end{aligned}
$$

Week Swap

Algorithm 1:

Input: A feasible solution to the opponent assignment model OPP.

Step 1. Calculate c_{ij} $\forall i, j$ and let min $= \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}^2$.

Step 2. Find i, j s.t. $c_{ij} \geq f$. If there are no such i, j, then end the algorithm.

Step 3. Find l, k s.t. $x_{ilk} = 1$ and $x_{jl(k+1)} = 1$. If there are no such l, k, then go back to Step 2.

Step 4. **for** m=1:n-1 **do**

Swap $x_{ij(k+1)}$ with x_{ijm} $\forall i, j$.

if $\sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}^2 < \min$

if new solution is feasible wrt PAT model then update the solution.

elseif reverse the swap.

end

Step 5. Go back to Step 3.

Description of the League

- Organized as a mirrored double round robin tournament.
- Consists of 18 teams and each team follows a home-away pattern with at most 1 break, with break-free first and last weeks.
- Hard constraints are the requirements with respect to home-away concerns, and soft constraints are in general related to resource scarcity and availability of teams due to international games.
- At the beginning of the season, fixture with only games and their weeks are announced.
- The exact days (which day of the week) each game will be played on is announced only one week before the game to avoid any game matching.

Scheduling of a half-season

- Canonical schedule (circle method) is used.
- Each team (except the randomly chosen team 17) follows the same sequence of opponents (teams 1-2-8-4-6-10-12-14-16-18-15-13-11-9-5-3-7), i.e., template forces each team to follow the same sequence of opponents.
- Odd numbered teams play home during the first week and alternate to away the next week.
- If a team encounters its own number in the sequence then it plays with team 17 and has a break on that week.
- After this initial assignment, the schedule is updated manually to cover for specific soft constraints for teams.

Scheduling of a half-season

- The canonical form results in maximizing the carry-over effects (Lambrechts et. al. 2016)

TABLE 2. Coe values of TPFL between 2012 and 2016.

Season	Coe
2012-13	3666
2013-14	3818
2014-15	3707
2015-16	3820
Average	3752.75

Heuristic results for the TPFL

- All computations are run via CPLEX 12.6 on a 3.40-GHz Intel Core i7-3770 Processor with 32.0 GB of RAM.
- The exact integer program for an 8-team league takes more than 48 hours to give a solution with a 90% gap.
- We first provide the computational results for 8, 10, 12, 14, 16 and 18 teams.

TABLE 3. Results of our heuristic.

Number of teams	Coe Value (1-Break)	Time (min)
8	104	0.20
10	192	0.44
12	318	2.73
14	446	2.73
16	626	6.30
18	944	23.90

Heuristic results for the TPFL

- For an 18 team, the initial coe (after pattern and opponent assignment) is equal to 1104.
- After week swap, we reduce it to 944 in about 24 minutes.
- 944 is a better coe value than many of the European leagues in Season 2015-16 (e.g., coe values for Germany and Spain are 1054 and 5183, respectively.)

When more than 1 break is allowed for a team

- Sticking to certain H and A pattern restrictions makes it hard to lower coe values.
- i) Minimize the number of consecutive breaks. For example, we would prefer AAHAA over AAAHA which both have 2 breaks.
- ii) Minimize the difference between the number of breaks for each team.

When more than 1 break is allowed for a team

- We adapt the Week Swap phase using the following model where we check for both feasibility and number of breaks.
- h_{ik} equals 1 if team i plays at home on week k, and 0 otherwise,
- r_{ik} equals 1 if there's a break on week k for team i, and 0 otherwise
- b_i represent the number of breaks of team i.
- Parameters u and v represent the lower and upper bounds on the number of breaks for each team to address concern ii) above.

When more than 1 break is allowed for a team

$$
\begin{aligned}
\text{BRK:} \quad & \text{Min} \sum_i b_i \\
\text{subject to} \quad & h_{ik} + h_{jk} = 1 && i, j | x_{ijk} = 1, \\
& h_{ik} + h_{i,k+1} - 1 \leq r_{ik} && i = 1, \ldots, n, \quad k = 1, \ldots, n-1, \\
& 1 - h_{ik} - h_{i,k+1} \leq r_{ik} && i = 1, \ldots, n, \quad k = 1, \ldots, n-1, \\
& \sum_k r_{ik} = b_i && i = 1, \ldots, n, \\
& u \leq b_i \leq v && i = 1, \ldots, n, \\
& 1 \leq h_{ik} + h_{i,k+1} + h_{i,k+2} \leq 2 && i = 1, \ldots, n, \quad k = 1, \ldots, n-1.
\end{aligned}
$$

When more than 1 break is allowed for a team

- When Algorithm 1 is run with BRK model in Step 4, with three different initial solutions.

TABLE 4. Coe values and number of breaks for the heuristic with the BRK model. ($n = 18$).

	Random		Canonical		TPFL	
Break (Each)	Coe	Break	Coe	Break	Coe	Break
$1 \leq x \leq 2$	650	31	698	29	580	32
$2 \leq x \leq 3$	544	45	502	42	466	45
$2 \leq x \leq 4$	544	45	502	42	466	45

When more than 1 break is allowed for a team

- When we relax constraints on the number of breaks (i.e., remove the lower bound from the BRK model):

TABLE 5. Coe values and number of breaks when break distribution among teams is not considered.

	10 Teams		12 Teams		14 Teams		16 Teams		18 Teams	
Break (Each)	Coe	Break	Coe	Break	Coe	Break	Coe	Break	Coe	Break
$x \leq 1$	192	8	318	10	446	12	626	14	944	16
$x \leq 2$	144	12	212	18	344	26	472	30	646	30
$x \leq 3$	144	12	212	16	302	24	396	34	556	40

When more than 1 break is allowed for a team

- How our heuristic performs under the break specifications of three real leagues which do not use the canonical schedule.
- Using the same number of individual breaks, our heuristic was able to create a schedule with smaller coe values in each case, with an average decrease of 18% in coe values.

TABLE 6. Comparison of our heuristic with three real league schedules for season 2015-16.

	Num. of teams	Coe	Breaks (Each)	Break	Coe (H)	Break (H)	Decrease (%)
Czech Rep.	16	742	$x = 1$	16	626	14	15.63
Germany	18	1056	$0 \leq x \leq 1$	16	944	16	10.60
France	20	719	$0 \leq x \leq 4$	45	530	56	26.28

Conclusion

- We addressed a novel and a relevant research problem.
- We focused on fairness considering carry-over effects and number of breaks at the same time.
- We proposed a secondary method that captures the trade-off between number of breaks and coe values.

Conclusion

THANK YOU, QUESTIONS?

Template

TABLE 7. Template for an 18-team league.

Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
1-11	18-1	1-8	12-1	1-14	2-1	1-17	4-1	1-5
6-2	2-7	16-2	2-9	3-2	8-3	15-2	2-11	18-2
3-14	12-3	3-13	10-3	7-4	4-6	3-4	5-3	3-17
17-4	4-5	9-4	4-16	6-5	5-9	16-5	7-6	15-4
5-10	13-6	17-5	5-7	9-8	17-7	6-8	8-18	6-9
7-13	8-17	6-12	17-6	15-10	10-11	9-7	17-9	16-7
16-8	10-9	7-10	8-15	11-13	12-15	18-10	10-16	11-8
9-12	11-15	11-18	14-11	18-12	13-16	11-12	12-14	14-10
15-18	14-16	15-14	13-18	16-17	14-18	14-13	13-15	12-13
Week 10	Week 11	Week 12	Week 13	Week 14	Week 15	Week 16	Week 17	
9-1	1-7	16-1	1-6	3-1	15-1	1-10	13-1	
2-13	14-2	2-12	10-2	2-8	2-5	17-2	2-4	
7-3	3-6	9-3	3-16	4-13	11-3	3-15	18-3	
4-18	11-4	4-14	12-4	5-12	10-4	4-8	8-5	
5-11	15-5	5-18	14-5	6-14	18-6	5-13	10-6	
6-16	12-8	6-15	18-7	7-15	8-7	6-11	14-7	
8-14	16-9	7-11	13-8	9-18	13-9	7-12	11-9	
10-12	13-10	8-10	15-9	17-10	12-16	9-14	12-17	
17-15	18-17	17-13	11-17	16-11	14-17	16-18	15-16	