Discrete Applied Mathematics 35 (1992) 301-312
North-Holland

Combinatorial aspects of construction of competition Dutch Professional Football Leagues

Jan A.M. Schreuder

Faculty of Applied Mathematics, University of Twente, Postbus 217, 7500 AE Enschede, Netherlands

Received 17 January 1990
Revised 1 August 1990

Abstract

Schreuder, J.A.M., Combinatorial aspects of construction of competition Dutch Professional Football Leagues, Discrete Applied Mathematics 35 (1992) 301-312.

Competitions are defined as a set of meetings between a number of clubs at preset dates under preset rules. Such a problem can be divided in two subproblems: firstly developing a Home-Away schedule with oriented egde-colourings of complete graphs and secondly assigning the clubs to the Home-Away patterns with a clustering algorithm.

Theoretical and real world aspects will be demonstrated by the construction of the Dutch Professional Football (US: soccer) Leagues.

Keywords. Timetable, schedule, combinatorics, sport.

Introduction

The KNVB (Royal Dutch Football Association) is responsible for the construction of the timetables for the professional football leagues in the Netherlands (see Table 1). In view of the increasing number of requirements which those timetables have to obey, the KNVB requested the Faculty of Applied Mathematics of the University of Twente to support them with the construction of the timetables (Huijbregts and Rijkhoek [6]). The research is carried out by Jan Schreuder (modelling and programming) and Jan Telgen (management aspects).

The timetable of the competition is not an isolated occurence. With the construc-

Table 1: Names clubs

Top-league		First-league	
PSV	PSV	Pec	PEC Zwolle
AJX	Ajax	Vvv	VVV
TWT	FC Twente	Vnd	Veendam
FEY	Feyenoord	Exc	Excelsior
RJC	Roda JC	Az	AZ
GNG	FC Groningen	Hrv	SC Heerenveen
BDB	BVV Den Bosch	Nac	NAC
FTS	Fortuna Sittard	Svv	SVV
VOL	FC Volendam	Gae	GA Eagles
HRL	Haarlem	Gfs	De Graafschap
RKC	RKC	Cam	Cambuur
SPA	Sparta R.	Hrc	SC Heracles'74
UTR	FC Utrecht	Ehv	Eindhoven
MVV	MVV	Ds9	DS'79
WII	Willem II	Tel	Telstar
VTS	Vitesse	Hel	Helmond Sport
DHG	FC Den Haag	Rbc	RBC
NEC	NEC	Emn	Emmen
		Wag	Wageningen

tion we have to take into account the requirements of the different parties involved like municipalities, police, railways, FIFA (International Football Federation), the clubs and press (especially television).

One of the starting points with the development of the competition timetables is that it could not be carried out by a once and for all established computerprogram. The reasons are that the construction of the timetable is only needed once a year and it has to be adjusted every year in view of the fast changes in the environment of football. A computerprogram in itself, however, can offer important support with the construction.

The main objective of our approach is that the final timetable to use is chosen based upon such norms and qualifications, that the interests of all parties involved can be handled in a balanced way by the KNVB.

In this paper we will discuss the construction of the Dutch Professional Football Leagues, as executed in 1989/90. The top-league consists of 18 clubs and 126 requirements (after selection!) have to be considered. The emphasis is on the way the combinatorial aspects are taken into account.

Problem

A timetable or roster consists of a set of meetings (or resources) assigned to time periods. The well-known term scheduling is, certainly in literature, reserved for assigning jobs/activities to machines.

A timetable of a sportscompetition consists of a set of fixed dates or rounds. In each round each club plays one match: either at home or away or is free (odd number of clubs). In such a competition all clubs meet each other twice: one home match and one away match: one match in the first half of the competition and the

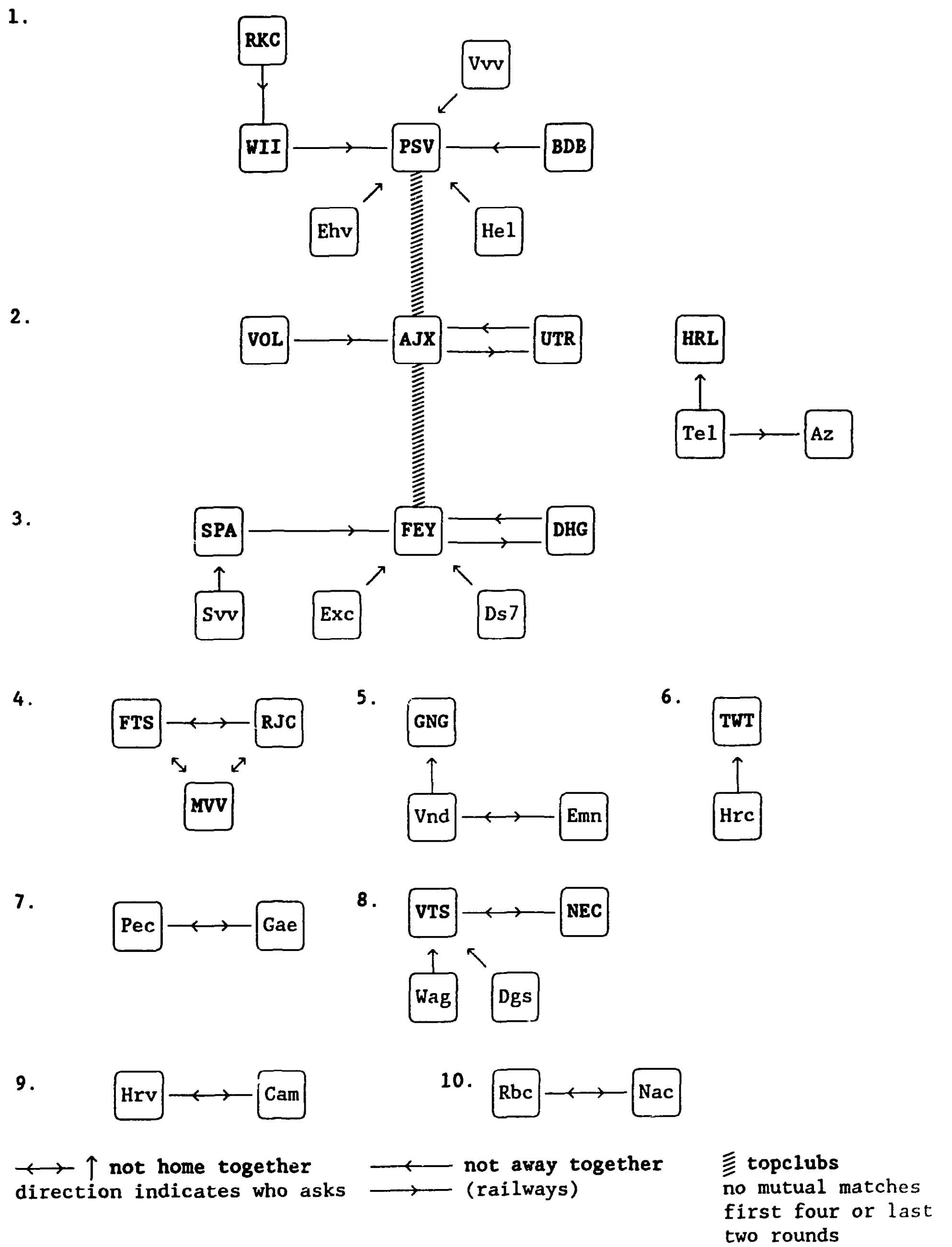


Fig. 1. Mutual relations clubs 1989/90.

other in the second half. Once the first half of a competition is scheduled, the second half is fixed and consists of the complement of the first half. Only rounds can be exchanged, not single matches [11, 15, 16] without disturbing the optimal alternating pattern of home and away matches for the clubs.

The dates for the matches are fixed by KNVB taking into account the already established international matches and possible European Cup matches. Based on these dates we have to decide per round which clubs would meet each other and which one of each pair plays at home.

The requirements of all parties involved are now-a-days so complex and often in conflict that all possible timetables - for n clubs $O(n \cdot (n-1)!)$ - have to be considered. These requirements can be divided in the following three categories.

(1) *Commercial aspects.* Clubs which are located so close together that they share the same fans, like to play their home matches in different rounds. Also, other events which attract the same kind of public should be avoided such that the club plays an away match on that date.

(2) *Sportive aspects.* A club promoted to the top league plays the first match at home. Schedules should have an optimal alternating ordering of the home and away matches for each club. Before each round, all clubs should have played the same number of matches.

(3) *Organisational.* One of the most important requirements nowadays is based upon the behaviour the hooligans under the fans, e.g. clubs with those fans are not allowed to play away matches during the week [6].

Examples of relations between the clubs are given in Fig. 1. The ordering of the requirements is not fixed and changes from year to year.

The basic requirement of a competition is that it must be fair, i.e., all the parties involved have the feeling the above requirements are met in a balanced way and their wishes are taken seriously.

Mathematical formulation and representation

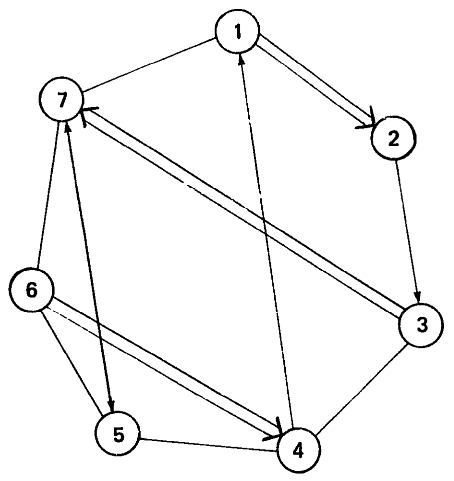
A mathematical formulation of the problem is the following [11].

For an even number $2n$ of clubs, let i and j denote the index for the clubs, and t $(=2n-1)$ denote the index for the rounds. Let x_{ijt} a zero-one variable with value one if club i plays at home against club j in round t, and value zero otherwise. Determine $X=[x_{ijt}]$ as a zero-one matrix with $n \cdot (2n-1)$ one's. Then, X is a competition if and only if

$$\sum_{i=1}^{2n} (x_{ijt}+x_{jit})=1 \qquad \forall j, t \tag{1}$$

and

$$\sum_{t=1}^{2n-1} (x_{ijt}+x_{jit})=1 \quad \forall j \neq i. \tag{2}$$



1^{st} ROUND	2^{nd} ROUND
1 ⇒ 2	2 → 3
7 ⇐ 3	1 ← 4
6 ⇒ 4	7 → 5
5 free	6 free

Hamiltonial path: 6⇒4→1⇒2→3⇒7→5

Fig. 2. Associated graph competition seven clubs.

(1) One of each pair per round plays at home.
(2) Each club plays one match against all opponents.

This formulation, however, cannot be used to solve the problem in real time. The value is to use it as a test for the correctness of any proposition for constructing a timetable (validation). Based upon this, the link between competitions and edge-colourings of complete graphs can be made (see Fig. 2). The vertices of such an associated graph represent the clubs, the edges represent the matches of the first half (or second) of the competition. The different rounds are represented by colouring the edges (one colour per round). The orientation of the edges (direction) points out which of the two clubs incident with such an edge plays at home. We used K_{2n-1}, the complete graph on $2n-1$ vertices, which represents the first half of the competition for an odd number of clubs.

Analysis

As already mentioned in the problem section, the construction of the timetable is NP-hard. Therefore we divide it into the following two subproblems in order to find an approximate solution as good as possible. This approach is illustrated for six clubs in Fig. 3.

Firstly, we develop a basis schedule in which the Home-Away pattern (HAP) is fixed. The HAP is represented in Fig. 3(a), where in each row i the opponents in the different rounds of a club i are given and a + sign denotes a home match for i. For example, $HR_{3,4} = -1$ means that the club which gets HAP_3 plays in the fourth round away against the club with HAP_1.

Secondly, we assign the clubs to the HAP's such that as much as possible requirements are fulfilled given their mutual weights. This assignment is represented in Fig. 3(b), where the capital letters denote the real clubs with their mutual relations and requirements to fulfill. In Fig. 3(c) the competition timetable is given as a combination of Fig. 3(a) and 3(b).

The basis schedule (HAP) is developed by determining an oriented edge-colouring of K_{2n-1}. The object is to find such a schedule that the number of breaks - two home matches or two away matches in succeeding rounds - is minimum.

The construction of a HAP is based on the following constructive theorem [12].

Theorem. *The HAP of an odd number of clubs contains no breaks. The HAP of an even number of clubs, however, contains* $2n-2$ *breaks.*

The following starting points are essential for the construction of the HAP's.

Start with an odd number of nodes: $2n-1$. Put these nodes in the form of a $(2n-1)$-gon (as on a circle with equal distances between neighbours). According to familiar topological properties, K_{2n-1} can be partitioned in $2n-1$ partial subgraphs G such that each G consists of one boundary edge and its parallels (see Fig. 2). Those graphs G have no edge in common. Also known is that K_{2n-1} can be coloured with $2n-1$ colours (chromatic number χ). If we assign to each G a different colour, then G represents the pairings of one round. The node which is not incident with an edge of G is called a free node and represents the club which has no match in that round.

In order to decide which one of each pair of nodes plays at home, all the edges have to be directed (orientation). This direction is carried out such that each two

	ROUND			+ home − away	
HAP	1	2	3	4	5
1	+2	−4	+6	+3	−5
2	−1	+3	−5	−6	+4
3	+5	−2	+4	−1	+6
4	+6	+1	−3	+5	−2
5	−3	−6	+2	−4	+1
6	−4	+5	−1	+2	−3

(a)

HAP		CLUB
1	↔	A
6	↔	B
4	↔	C
5	↔	D
2	↔	E
3	↔	F

(b)

	DATE			+ home − away	
CLUB	1	2	3	4	5
A	+E	−C	+B	+F	−D
B	−C	+D	−A	+E	−F
C	+B	+A	−F	+D	−E
D	−F	B	+E	−C	+A
E	−A	+F	−D	−B	+C
F	+D	−E	+C	−A	+B
	6	7	8	9	10

(c)

Fig. 3. (a) A basis schedule for six clubs. (b) Assignment in view of requirements. (c) Competition timetable combination of (a) and (b).

succeeding rounds (G_i and G_{i+1}) form a directed Hamiltonian path. This path consists alternatingly of the first and second round (the graph which consists of G_i and G_{i+1} is bipartite).

All the nodes have one in-going and one out-going edge (or arrow) except the start node (out-going) and the end node (in-going). When an in-going edge represents a home match (the out-going an away match), then there are no breaks in the second round G_{i+1}. So, the HAP of an odd number of breaks contains no breaks.

For an even number of clubs, add a $2n$th club which plays against the free node in each round. Give this $2n$th club a HAP with no breaks. As each original free node has in-going edge in the proceeding round and an out-going in the succeeding round of the round when he is not playing, there are exactly $2n-2$ breaks (no break in the first round).

Executing the construction in the above described way needs no backtracking. Each HAP has a complement [16].

Properties of the HAP as applied to the competition timetable are given in Fig. 4.

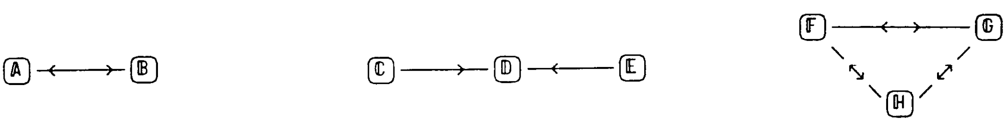
Important is that once the HAP's are established, they may not be changed. This is due to the fact that the different leagues are related through their HAP's. A club located in the neighbourhood of a club in a higher league wants to play a complementary HAP (see Fig. 4(a)) like HAP_1 and HAP_2 in Fig. 3(a).

Assigning the clubs to HAP's is a NP-hard problem based on the conflicting objectives and the uncertainty of the mutual weights for the different requirements. In case of n clubs, the solution space of $n!$ has to be evaluated.

Models

There are a lot of models available with which the assignment of clubs to given HAP's can be formulated.

A first formulation only allows requirements for individual clubs. Suppose, there are n clubs and, consequently, n HAP's. Let i denote the index for the clubs and j denote the index for the HAP's. Let x_{ij} be a zero-one variable with value one if club i gets HAP_j and value zero otherwise. Let c_{ij} be a weight function indicating the relative value of the assignment. For example, a club wants to play not at home at a certain date, then only some HAP's are allowed (see − sign in Fig. 3(a)) and



(a) (b) (c)

Fig. 4. (a) A and B no home matches together. (b) C and E no home matches together with D. (c) F, G and H no home matches together.

the importance is not the same for each club or date. Then, a competition timetable can be constructed with the following model.

$$\text{Maximise} \quad \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij}, \tag{3}$$

$$\text{subject to} \quad \sum_{i=1}^{n} x_{ij} = 1 \quad \forall j, \tag{4}$$

$$\sum_{j=1}^{n} x_{ij} = 1 \quad \forall i. \tag{5}$$

Of course, each HAP should be assigned to exactly one club (4) and each club gets one HAP (5).

This model would be sufficient and easy to solve, if the requirements could be restricted to preferences of clubs to HAP's. But, quite common, combinations of clubs (see also Fig. 1) and HAP's are important. For example, two closely located clubs require complementary HAP's. The model should be extended in the following way.

Let i and j denote the index for the clubs, p and q denote the index for the HAP's. Let h_{pq} denote the number of home matches together if HAP_p and HAP_q are combined and x_{ip} a zero-one variable with value one if club i is assigned to HAP_p. Then the following linear and quadratic model can be constructed.

$$\text{Maximise} \quad \sum_{i} \sum_{p} c_{ip} x_{ip} + \sum_{i} \sum_{j} \sum_{p} \sum_{q} h_{pq} x_{ip} x_{jq}, \tag{6}$$

$$\text{subject to} \quad \sum_{i} x_{ip} = 1 \qquad \forall p, \tag{7}$$

$$\sum_{p} x_{ip} = 1 \qquad \forall i, \tag{8}$$

	CLUSTER F			GM			FA			
HAP	c	c	c	c	c	c	c	c	c	← weights
1				1						⇒ one club
2	1				1		1	1		
3						1	1			
4		1				1		1	1	
5			1	1						
6					1				1	
Club										
F	1	1	1				1	1	1	⇒ one HAP
G				1	1	1				
M				1	1	1				
A							1	1	1	

MIN $\underline{c}\,\underline{x}$

S.T. $A\underline{x} = \underline{1}$

Fig. 5. Set partitioning.

$$\sum_j \sum_q h_{pq} x_{ip} x_{jq} \leq hm_{ij} \quad \forall i, p. \tag{9}$$

Like in the preceding model (7) and (8) assure: one club ⇔ one HAP. If hm_{ij} denotes the maximal number of allowed home matches together of club i and j, then (9) assures that this number is not violated.

Known from the literature (quadratic assignment [10]) is that these kinds of models, even for small sizes ($n \leq 17$), are hard to solve.

A nice and illustrative example for modelling all possible requirements is the set partitioning approach as presented in Fig. 5.

Each club or combination of clubs (= cluster) is represented in the columns and the HAP's in the rows. We put one in a column if a cluster uses the corresponding HAP. Of course, each scheme may only be used by one club. In order to assure that each club uses only one scheme we add a row for each club putting a one if the club belongs to the cluster of that column. The order in which the requirements are fulfilled determines the composition of the clusters.

In order to solve the described problem, two approaches in literature are worth mentioning. A representative for the level-by-level heuristics is branch-and-bound with columngeneration [9] and for the stochastic search methods simulated annealing [7, 8]. For our final approach we took the properties which uses the structural and feasible part from the level-by-level methods and the higher chance for finding a global solution from the stochastic search ones.

Approach

As pointed out in literature [3], the kind of problems as described here cannot be solved in real time with exact algorithms. Therefore, we settled for the following approach, see Fig. 6.

The requirements are divided into hard ones – the constraints – and soft ones – the objective function – [2]. The requirements in between are alternately classified as hard or soft ones in order to determine the robustness of the found solutions. This division is prescribed by a national committee (commissie Waal) in which the municipalities, the police and the railways are represented.

Important here is that the classical concept of feasibility is exchanged for acceptable [14]. This means that the way of approaching the problem convinces people that the solution offered is the best one considering the circumstances. Feasibility viewed as how the requirements are fulfilled – some will be violated ! – is less important in real world applications.

In order to reach for a solution, we firstly generate all possible combinations in an implicit way. We keep only those combinations which fulfill the hard demands. Of course, this generation have to take place in a constructive way. We actually built our solution space starting with partial solutions without violating the optimal solution. We determine these solutions with the clusters which are described for the set-partitioning approach in the model section. This (heuristic) way of approach is called construction and partitioning [17].

The possibility to combine clusters in order to determine a solution while reducing the solution space, depends strongly on whether the clusters are disjunct or not.

(A) No clubs in common: $\bigcap = \emptyset$.
Reduction if clubs have at least one HAP in common.

(B) Clubs in common: $\bigcap \neq \emptyset$.
Reduction if
(i) common clubs have not the same HAP;
(ii) not-common clubs like (A).

These rules and even more sophisticated are known for the satisfiability problem [4, 13].

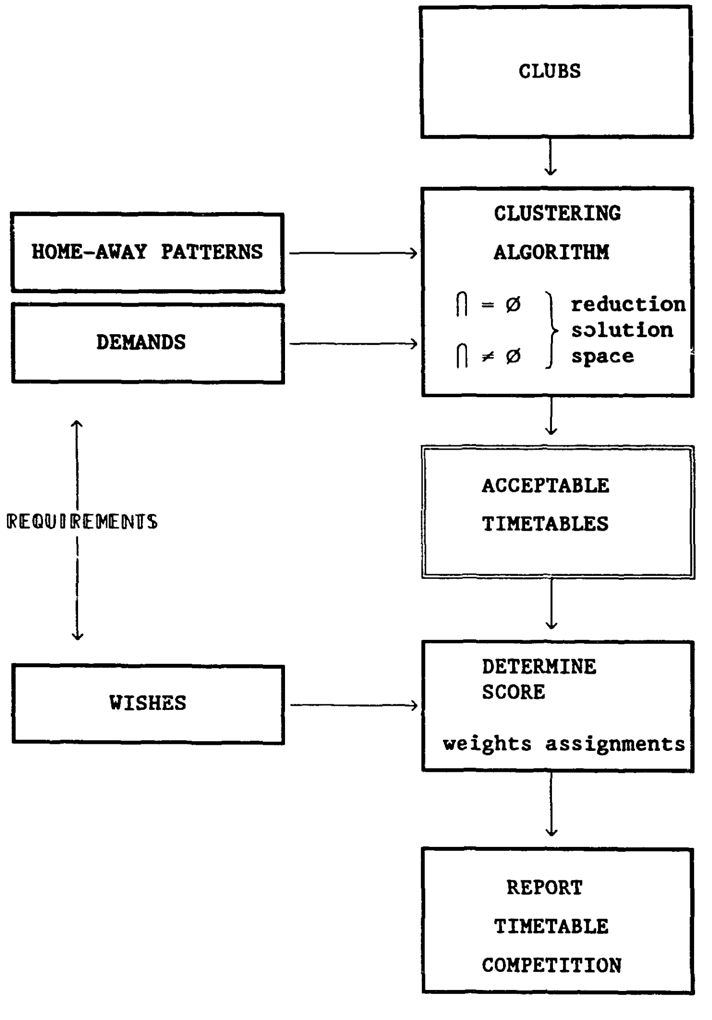


Fig. 6. Applied approach.

The solution space for the example of Fig. 5 exists of cluster FAGM with respectively the HAP's 2315, 2415 and 4615.

Secondly we give a weight to all wishes in order of their importance. Of course, the mutual weight of the wishes is hard to establish. Not only is the given preference not unique (transitivity does not hold), but also the combination realized in the final solution determines the value of the wishes.

Finally we calculate the score of each remaining combination (just keep the 5 highest) and the frequency distribution of the wishes fulfilled. These results are presented to the football leagues.

The approach described in this paper is applied to the construction of the timetable for the competition 1989/90 of the top-league in Holland consisting of 18 clubs. A total of 126 requirements are considered of which 80% is realized. For this 4700 timetables are constructed (the reduced solution space) and then evaluated on an Olivetti-M24 PC (≤ 2 minutes). Of course, this construction is repeated several times in order to get some insight in the sensibility of the weights.

Conclusion

The construction of timetables for the Dutch Professional Football Leagues leads to an interesting problem with different theoretical and practical aspects which are strongly related. In view of fast changing circumstances of our society there are still a lot of aspects to be considered.

More theoretical based is the ordering of the timetable, minimal lengths of same opponents, and equal opportunities for all clubs in the solution space of the demands. More practical is the division of the requirements in demands and wishes, and the weight of those wishes. Also the significance of recent developments like neural networks, simulated annealing, tabu search [5] and genetic generation could be taken into account. However, there is not always time and room for an extensive search for those solutions. What really counts is the way of approach, not the used techniques.

Computerprograms can play an important part in realizing and supporting solutions.

References

[1] D.C. Blest and D.G. Fitzgerald, Scheduling sports competitions with a given distribution of times, Discrete Appl. Math. 22 (1988) 9–19.

[2] H.A. Eiselt and G. Laporte, Combinatorial optimization problems with soft and hard requirements, J. Oper. Res. Soc. 38 (1987) 785–795.

[3] H.A. Fleuren, A computational study of the set partitioning approach for vehicle routing and scheduling problems, Ph.D. Thesis, University of Twente, Enschede (1988).

[4] M.R. Garey and D.S. Johnson, Computers and Intractability (Freeman, San Francisco, CA, 1979).

[5] A. Hertz, Tabu search for large scale timetabling problems, Rept. O.R.W.P.89/4, École Polytechnique Fédérale de Lausanne, Lausanne (198[illegible]).
[6] J.G. Huijbregts and B. Rijkhoek, Secretariaat Betaald Voetbal KNVB, 1989.
[7] S. Kirkpatrick, C.D. Gelatt Jr and M.P. Vecchi, Optimization by simulated annealing, Science 220 (4598) (1983) 671–680.
[8] R. Kuik and M. Salomon, Multi-level lot-sizing problem: evaluation of a simulated-annealing heuristic, European J. Oper. Res. 45 (1990) 25–37.
[9] C.C. Ribeiro, M. Minoux and M.C. Penna, An optimal column-generating-with-ranking algorithm for very large scale set partitioning problems in traffic assignment, European J. Oper. Res. 41 (1989) 232–239.
[10] C. Roucairol, A parallel branch and bound algorithm for the quadratic assignment problem, Discrete Appl. Math. 18 (1987) 211–225.
[11] J.A.M. Schreuder, Constructing timetables for sport competitions, Math. Programming Stud. 13 (1980) 58–67.
[12] J.A.M. Schreuder, Timetables for sport (soccer) competitions, Memorandum Nr. 349, TH Twente-TW, Enschede (1981).
[13] J.A.M. Schreuder, Application of a location model to fire stations in Rotterdam, European J. Oper. Res. 6 (1981) 212–219.
[14] J.A.M. Schreuder and J.A. van der Velde, Timetables in Dutch high schools, in: Operational Research '84 (Elsevier, Amsterdam, 1984) 601–612.
[15] D. de Werra, Scheduling in sports, Discrete Appl. Math. 2 (1980) 327–337.
[16] D. de Werra, Some models of graphs for scheduling sports competitions, Discrete Appl. Math. 21 (1988) 47–65.
[17] S.H. Zanakis, J.R. Evans and A.A. Vazacopoulos, Heuristic methods and applications: a categorized survey, European J. Oper. Res. 43 (1989) 88–110.