Lancaster University Management School
Working Paper
2004/034

A Rich Model For Scheduling Umpires For An Amateur Cricket League

Wright, Mike

The Department of Management Science
Lancaster University Management School
Lancaster LA1 4YX
UK

A Rich Model For Scheduling Umpires For An Amateur Cricket League

by

M.B. Wright

Department of Management Science

Lancaster University

LA1 4YX

U.K.

tel. +44 1524 593846

Fax +44 1524 844885

email m.wright@lancaster.ac.uk

Abstract

This paper describes a problem faced every year by the Devon Cricket League in England. Every league match requires two officiating umpires. There are various rules relating to the deployment of umpires, and rather more preferences to be considered which fall short of being definite constraints.

The requirement was to produce a computer system which would allocate umpires in a satisfactory manner without the need for human interaction, since the users of the system will be unqualified amateurs. This necessitated the construction of a "rich" model incorporating several solution attributes of various types.

This paper describes this model in detail, together with the solution procedure, a form of metaheuristic search known as subcost-guided simulated annealing.

The system has been put into successful use for the 2003 and 2004 seasons and it is expected that its use will continue indefinitely.

Key Words – Scheduling, timetabling, sport, cricket, rich models, metaheuristics, simulated annealing, multiple objectives, subcost guided search

Scheduling of sports officials

There have been many papers in the academic literature concerned with the scheduling of sports fixtures. Some of these have been theoretical in nature; recent examples include Easton *et al*. (2003), Urban and Russell (2003), Trick (2003) and Schonberger *et al.* (2004). Others have focused upon case studies, for example Willis and Terrill (1994), Nemhauser and Trick (1998), della Croce *et al.* (1999) and Wright (2004).

However, papers concerned with the scheduling of sports officials are very few and far between. Apart from Evans et al. (1984), Evans (1988) and Wright (1991) there appears to be nothing. This is surprising given that it must be an important issue for all sports leagues and tournaments world-wide.

This paper concerns the scheduling of umpires for the Devon Cricket League, an amateur league in England. The first part of the paper describes the problem in detail and the way in which it was modelled, using a "rich" model. The second part then describes the solution method and the results achieved.

Rich models

A new class of model has recently been defined – a "rich" model. These are models which are sufficiently detailed that they can produce satisfactory solutions to real complex problems without the need for human interaction either during or after the solution procedure. This is of crucial importance when the users of such models are either unwilling or unqualified to interact with a

system in a fruitful way, or when a solution derived from an insufficiently detailed model would not even provide a useful starting point for such interaction.

The concept of a rich models is not in itself new. Many complicated problems have been approached in this way for many years. However, the speed and power of modern computers is now such that it should be possible to tackle a very wide range of problems in this way.

Rich models are due to be the topic of a special issue of the *European Journal of Operational Research* in late 2004.

The Devon Cricket League

The Devon Cricket League is an amateur league which spans one of the largest counties in England. Matches occur on Saturday afternoons from late April or early May until late August or early September.

The League is currently divided into eight divisions. The top division is called the Premier Division, followed by Divisions A, B, C, D West, D East, E West and E East. The two D divisions are of equal status but geographically separated; likewise the two E Divisions. Each division contains ten amateur clubs (or about ten in the case of the lower divisions). Every club has its own home ground and plays fixtures against every other club in its division exactly twice, once at home and once away. The fixtures are arranged so that one of these matches takes place in the first half of the season, and one in the second half. At the end of each season clubs are promoted or relegated in the usual way.

The umpire allocation exercise described here involves only the top three divisions; umpires for the other divisions are sorted out later on a more local basis. Every club in the top three divisions has a match every Saturday during the season, which lasts eighteen weeks.

Each match requires two umpires who are paid a small fee as well as travel expenses. As the league is run on a financial shoestring, it is of vital importance to keep these expenses low; however, there are also several other considerations to take into account when deciding which umpires should be allocated to each match. Thus the problem of allocating umpires to matches after they have been timetabled is decidedly non-trivial.

This task falls to the league secretary – an unpaid role usually occupied by a retired cricket player who is unlikely to be a sophisticated computer user. The current occupant of this role therefore appealed for help in allocating the League's umpires.

The umpire allocations are derived in two separate tranches, one for the first half of the season and one for the second. The timing of these tranches is under the control of the league secretary. The first tranche needs to take place a few weeks in advance of the start of the season. The timing of the second tranche depends on two conflicting factors; the need to give advance information to umpires and the desirability of taking account of events during the first half of the season.

Information required

As the model is an rich one, it requires a fair amount of detailed information before it can be used. Information is required regarding the clubs, the matches, the umpires and the rules, all of which must be in a format that can conveniently be entered and amended by the system user.

For each club, the data to be input is its name, its address (in the form of geographical co-ordinates) and the division in which it is playing.

For each match, the system requires the home team, the away team and the date.

Umpire information is a little more complex. The league has a pool of around fifty umpires, who all live within the county of Devon or in a neighbouring county (Cornwall, Dorset or Somerset), near to one of Devon's borders. Their names and home addresses (again in co-ordinate form) must all be entered.

Many umpires are not available every Saturday of the season; indeed, some may only be available for the occasional date. Therefore full availability information must be entered.

Not all umpires are equal. Each umpire is allocated a rating (or status) of 1, 2, 3 or 4 by the league, reflecting his quality and experience. This affects the matches at which he (they are all men) is allowed to officiate (see later). For each umpire the user also decides upon a target number of matches in each division, in each half of the season.

A variety of types of rule can be entered. One essential set of rules defines the effect of the umpire status information. The simpler form of this rule simply states that umpires of status x cannot be used for matches in division z. In its more complex form, the rule says that a pair of umpires with statuses $\{x, y\}$ cannot officiate at a match in division z.

Another type of rule says that umpire x must, or must not, be allocated to a specific match m. Alternatively it may be specified that an umpire must have a match (without specifying which one) on a given date. Slightly more complex is a rule that says that matches m and m' must not have an umpire in common. In cases where m and m' both involve the same pair of teams, this rule is automatic and need not be specified explicitly.

Other rules concern maxima and minima for incidences of specific umpires with specific clubs, with specific clubs at home, and with specific other umpires. In the first case, the most likely maximum is zero, when an umpire has a particular association with a club, perhaps as an ex-player. In the second case, for example, an umpire may have made a special request to have, or not to have, a match at a particular ground at some time during the season. In the third case, there may be pairs of umpires who find it difficult to work together, or who are required to be deliberately paired together a certain number of times for training purposes (where one umpire is highly experienced and the other is new).

As only half of a season is being solved in a run, the rules regarding maxima and minima (which are taken to apply to the whole season) need to be temporarily changed during the program runs. The maxima and minima are halved, with minima rounded up and maxima rounded down for the first half of the season, but with minima rounded down and maxima rounded up for the second half

of the season. This is done because in other ways scheduling the second half is more challenging, as notice must be taken of allocations in the first half.

For the second half of the season, the first half's allocations are used as input, but should be changed manually before running the system to reflect any emergency changes that were made (e.g. because of illness). It may also be desirable to change one or more rules, and the targets and availability details for the second half of the season can also be changed at this stage.

Appendix 1 gives samples of input information used for scheduling the second half of the 2003 season for the Devon Cricket League.

Objectives

As well as the rules, the problem involves a number of objectives or preferences. Although these are not constraints, they must all be modelled to ensure that the system can produce solutions of high quality that will be immediately acceptable to the league secretary.

These preferences are modelled by means of subcosts or penalties within the objective function. The formulations and relative weightings of these subcosts were chosen after detailed discussions and a certain amount of trial and error. Inevitably the formulations and weightings are to a large extent arbitrary and cannot be regarded as absolute.

When the second half of the season is being scheduled, the costs refer to the whole season's schedule, as there needs to be interaction between the two halves of the season when calculating

costs. For example, for umpire X to be allocated to ground Y on the first date of the second half of the season will be a decidedly bad thing if X was also at Y on the last date of the first half of the season.

The various costs are as follows. A number in **bold** shows the present value of a parameter that can be varied by the system user.

"Rule-breaking" costs

When user-input rules are broken, this is not regarded as leading to an infeasibility, but to a high penalty cost. These penalties are as follows. They are all to a large extent arbitrary; they must be large enough to ensure that rules are not broken unless they have to be, but not so large that infeasible solutions cannot be used during the search, in order to lead to better feasible solutions. Thus a certain amount of trial and error has contributed towards their determination.

No umpire allocated to a match – 4000 units

Only one umpire allocated to a match – 2000

Umpire not allocated to a match to which he must be allocated – 1000

Umpire allocated to a match to which he must not be allocated – 1000

Two matches have the same umpire when this is not allowed – 750

Umpire has no match on a date when he must have one – 500

Umpire (or combination of umpires) of insufficient status – 250

Umpire too often at ground – 150 times the square of the excess

Umpire not often enough at ground – 150 times the square of the shortfall

Umpire too often with club – 100 times the square of the excess

Umpire not often enough with club – 100 times the square of the shortfall

Two umpires too often together – 50 times the square of the excess

Two umpires not often enough together – 50 times the square of the shortfall

"Target-missing" costs

If an umpire of status S misses his target number of matches in division D by an amount Z, the cost is f(S) * g(D) * Z^2, multiplied by 10 if the target is zero (since a zero is probably a more definite request than any positive number).

Currently f(1) = **4**, f(2) = **2**, f(3) = **1** and f(4) = **4**, since umpires of status 4 are reserves, only to be used if there is no alternative, whereas for other umpires it is considered more important for the higher-status umpires to hit their targets.

Currently g(0) = **8** (where division 0 is shorthand for all divisions added together), g(1) = **10** (Premier Division), g(2) = **6** (Division A) and g(3) = **4** (Division B). This is because the targets are considered more important for the total number of matches and for the Premier Division matches, but slightly less important for the lower divisions.

Travel costs

The clubs' grounds and umpires' homes are all located within a 110 kilometre square area. Precise locations are represented using a 10 kilometre square grid within this area. Thus the west-east co-ordinate varies between 1 and 11 and the south-north co-ordinate likewise varies between 1 and 11.

Distances are calculated using Pythagorean distances. Thus, for example, the travel distance for an umpire living in square [2,9] to officiate at a ground located at [8,4] is approximated as $\sqrt{((8-2)^2 + (9-4)^2)} = \sqrt{61} = 7.8$ units (i.e. 78 kilometres).

As an important consideration is not just the total travel distance but the number of long journeys any single umpire must undertake, a definition of a "long" journey is required. This is currently set at **5.5** units (55 kilometres).

It is also common practice, if the two umpires allocated to a match both have to travel a long way, for one of them to be encouraged to give the other a lift if this does not involve too large a detour. Thus, within the computer system, a shared journey is defined as one where both umpires live a long way (i.e. more than **5.5** units) away from the match location and where the umpire giving the lift will not by doing so increase his journey length by more than **25**%.

The travel cost for an umpire is thus initially calculated as the total distance travelled (but with any shared journey only counting half) multiplied by a parameter **0.05**. If the number of long journeys for this umpire (with any shared journey counting half) is greater than 2 (if just the first half of the season is being scheduled) or 4 (if the second half is being scheduled) then this cost is multiplied

by {the excess plus one}. This is done to try to ensure that the long journeys are shared out in a reasonably equitable way between umpires, though inevitably those umpires living on the fringes of the county will have to travel more than those living near the centre.

Other costs

If an umpire and club are together more than once (first half of season being scheduled) or twice (second half being scheduled), the cost is **0.1** times the square of the excess. If an umpire and club are together twice with W intervening weeks, and $W < 6$, then there is a cost of **0.3** * $(6-W)^2$.

If an umpire and ground are together more than once, the cost is **0.2** times the square of the excess. If an umpire and ground are together twice with W intervening weeks, and $W < 8$, then there is a cost of **0.5** * $(8-W)^2$.

If two umpires are together more than once, the cost is **0.4** times the square of the excess. If two umpires are together twice with W intervening weeks, and $W < 8$, then there is a cost of **0.5** * $(8-W)^2$.

The above costs are required in order to spread out umpires between clubs, between grounds and among each other. Where there has to be a repetition of a pairing, it is preferable for these not to be too close in time. Squared formulations are used here because, for example, it is considered preferable for two umpires each to have an excess of one than for one umpire to have an excess of two.

If two umpires are together in a match for which they are both of a higher status than necessary, there is a cost of **0.1**. This is because, where a high-ranking umpire is used for a low-ranking match, part of the purpose is so that he may help his more junior co-umpire.

If the system is being run in order to amend an already existing schedule, another optional cost (of whatever size the system user requires – this is the only example of interactive input) may be specified for each change made to the input schedule.

Solution method

The solution approach used is subcost-guided simulated annealing (SGSA). This is the same as standard simulated annealing (SA), except that the cost increase is adjusted before being used in the acceptance criterion. The acceptance criterion is that a worsening perturbation is accepted if $R < e^{-(C'/T)}$, where R is a random number between 0 and 1, T is the temperature and C' is an amended cost increase, defined as $C' = Ce^{-\theta B/C}$, where C is the overall cost increase, B is the best decrease for any individual subcost and θ is a parameter. See Wright (2001a, 2001b) for a fuller explanation of SGSA.

SGSA requires four parameters – the start temperature, end temperature, number of iterations and θ. These are normally hidden parameters, as it is not expected that the user would understand their effect. Currently they are set at 4, 0.04, 5 million and 2.5 respectively, as a result of a fair amount of experimentation. The time taken to run the system with 5 million iterations obviously depends on the computer used, but can be expected to be of the order of ten minutes.

Results

The system produces various sets of output:

- A list of all matches, ordered by division and date, with names of umpires added
- A chronological list of matches, dates and partner umpires for each umpire, together with suggestions where it might be helpful for the umpires to travel together
- a table summarising the number of matches in each division for each umpire
- a table summarising the activity assigned to each umpire on each date
- a matrix showing umpire-club incidence
- a matrix showing umpire-ground incidence
- a matrix showing umpire-umpire incidence

See Appendix 2 for samples of output produced by the system when producing the allocations for the second half of the 2003 season.

At the time of writing, the system has been used for the Devon Cricket League for both halves of the 2003 season and the first half of 2004. The Devon Cricket League secretary has written:

> "We now have a more than satisfactory set of appointments. As the only people within the Devon League that knew of the programme's existence, the subcommittee casually asked various umpires what they thought of the season 2003 appointments. To our delight they thought it was the best list produced for many years as they were getting a fairer selection of games and a wider variety. Those going for full membership were all happy, as they had got what they had requested to enable them to apply for their full membership of the ACU&S (*the Association of*

Cricket Umpires and Scorers). The outcome of all this was that not only were the appointed umpires in the right place at the right time but were getting a fairer distribution of games as well as complying with League and ECB requirements. We now have umpires requesting to join the panel. To our subcommittee this has saved us many man hours of work, thus releasing us for other tasks which we were previously unable to do."

Implications for other research

This case-study shows that a highly satisfactory computer scheduling system can be produced for a complex problem, with detailed and fully implementable solutions being produced, even when the system user does not wish (or is not qualified) to have any interaction with the running of the system, as long as sufficiently rich models are specified and used. The approach used here could be used for a very wide variety of types of real-world application.

References

della Croce, F., Tadei, R. and Asioli, P.S. (1999) Scheduling a round robin tennis tournament under courts and players availability constraints. *Annals of Operations Research* 92, 349-361.

Easton, K, Nemhauser, G.L. and Trick, M.A. (2003) Solving the travelling tournament problem: a combined integer programming and constraint programming approach. *Lecture Notes in Computer Science* 2740, 10-109.

Evans, J.R., Hebert, J.E. and Deckro, R.F. (1984) Play ball – the scheduling of sports officials. *Perspectives in Computing* 4 (1), 18-29.

Evans, J.R. (1988) A microcomputer-based decision support system for scheduling umpires in the American baseball league. *Interfaces* 18 (6), 42-51.

Nemhauser, G.L. and Trick, M.A. (1998) Scheduling a major college basketball conference. *Operations research* 46 (1), 1-8.

Schonberger, J., Mattfeld, D.C. and Kopfer, H. (2004) Memetic algorithm timetabling for non-commercial sport leagues. *European Journal of Operational Research* 153 (1), 102-116.

Trick, M.A. (2003) Integer and constraint programming approaches for round-robin tournament scheduling. *Lecture Notes in Computer Science* 2740, 63-77.

Urban, T.L. and Russell, R.A. (2003) Scheduling sports competitions on multiple venues. *European Journal of Operational Research* 148 (2), 302-311.

Willis, R.J. and Terrill, B.J. (1994) Scheduling the Australian state cricket season using simulated annealing. *Journal of the Operational Research Society* 45 (3), 276-280.

Wright, M.B. (1991) Scheduling English cricket umpires. *Journal of the Operational Research Society* 42 (6), 447-452.

Wright, M.B. (2001a) Subcost-Guided Search – experiments with timetabling problems. *Journal of Heuristics,* 2001, 7(3), 251-260.

Wright, M.B. (2001b) Subcost-Guided Simulated Annealing. In: Ribeiro CC, Hansen P (Eds.). Essays and surveys in metaheuristics. Boston: Kluwer Academic Publishers, 631-639.

Wright, M.B. (2004) Scheduling fixtures for Basketball New Zealand. Lancaster University Management School Working Paper LUMSWP2003/039.

Appendix 1 – Samples of input data

Clubs

```
** ***************** ** **  *
 1 Abbotskerswell    6  4  P
 2 Barton            7  4  P
 3 Bovey Tracey      6  5  P
 4 Exeter            7  7  P
.............................
```

The two numbers are the co-ordinates of the home grounds of each club. The letter P denotes Premier Division.

Matches

```
                   Devon Cricket League 2003
                         Premier Division
Sat  3rd May   Abbotskerswell   v  Sidmouth
               Barton           v  Paignton
               North Devon      v  Exeter
               Plympton         v  Bovey Tracey
               Torquay          v  Sandford
Sat 10th May   Bovey Tracey     v  Abbotskerswell
      .............................

                         "A" Division
Sat  3rd May   Alphington       v  Axminster
      .............................

                         "B" Division
Sat  3rd May   Babbacombe       v  Clyst St George
      .............................
```

Umpires

```
** ***************** ** **  *  *********/********** ** ** * ** **
 1 J Anning          9  7  1           /X            5  2  1     5  2  0
 2 S Bilverstone     6  3  1           /             2  3  2     3  2  0
 3 G Curson          2  5  1  X  X  X  /  XXXX       4  1  1     3  1  0
      .............................
```

The first two numbers are co-ordinates of that umpire's home. The next number is the status of the umpire. The next nine fields refer to the first nine dates of the season: X if unavailable, blank if available. After the slash, the next nine fields are the same for the last nine dates. The next three numbers are the target number of matches in each division for the first half of the season; the next three numbers are the targets for the second half of the season.

Rules

```
Not ump/div  3 P
Not ump/div  2 P
Not ump/div  3 A
```

These lines mean that an umpire of status 3 should not have Premier Division matches; an umpire of status 2 should not have Premier Division matches; and that an umpire of status 3 should not have Division A matches.

```
Max ump/club J Anning         Sidmouth         0
      ..........................
```

Umpire Anning should not have any matches involving Sidmouth (because he used to play for Sidmouth).

```
Min ump/ump  B Perry          E Daniels        3
      ..........................
```

Umpires Perry and Daniels should be together at least three times during the season. This is because, to satisfy qualification requirements for the ACU&S (Association of Cricket Umpires and Scorers), an umpire must have three matches in a season with a specific current member of the ACU&S.

Appendix 2 – Output

APPOINTMENTS FOR YEAR 2003

Premier Division

Sat 3rd May	Abbotskerswell	v	Sidmouth	G Ripley		B Warren
	Barton	v	Paignton	E Daniels		J Hansford
	North Devon	v	Exeter	#J Anning		J Neville
	Plympton	v	Bovey Tracey	#J Harris		C Shelton
	Torquay	v	Sandford	D Moseby		B Perry
Sat 10th May	Bovey Tracey	v	Abbotskerswell	J Anning		L Varney

.....

"A" Division

Sat 3rd May	Alphington	v	Axminster	T Waldock		D Hodge

.....

"B" Division

Sat 3rd May	Babbacombe	v	Clyst St George	#V Gainey		K Jeffery

.....

The # sign denotes a "long" journey.

Schedule for umpire J Anning

Sat 3rd May	North Devon	v	Exeter	P	#	J Neville	
Sat 10th May	Bovey Tracey	v	Abbotskerswell	P		L Varney	
Sat 17th May	Exmouth	v	Seaton	A		B Bartlett	
Sat 24th May	Plympton	v	Barton	P	#	D Hawke	
Sat 31st May	Shobrooke Park	v	Tavistock	B		A White	
Sat 14th June	Budleigh Saltn	v	Plymstock	A		K Jeffery	
Sat 21st June	Paignton	v	North Devon	P		J Hansford	
Sat 28th June	Torquay	v	Exeter	P		B Warren	
Sat 12th July	Alphington	v	South Devon	A		S Bilverstone	
Sat 19th July	Plympton	v	Abbotskerswell	P	#	J Harris	TT
Sat 2nd Aug	North Devon	v	Barton	P	#	M Robertson	TT
Sat 9th Aug	Plymouth	v	Exmouth	A	#	J Raphael	
Sat 16th Aug	Exeter	v	Paignton	P		G Curson	
Sat 23rd Aug	Abbotskerswell	v	Sandford	P		B Perry	
Sat 30th Aug	Bovey Tracey	v	North Devon	P		T Waldock	

.....

TT stands for "Travel Together" – this is what the system regards as a potentially shared journey.

TYPES OF MATCH FOR EACH UMPIRE (second half of season only)

	Total	P	A	B
J Anning	7	5	2	0
S Bilverstone	5	3	2	0
G Curson	4	3	1	0
E Daniels	7	4	2	1

Number of matches in each division for each umpire

CLUB/UMPIRE INCIDENCE (full season)

	ABCDDHHJLMNPRRRSSVVWWBHEGGHHJMMRRSTCCFKLPRGHMPTWMHCQSSSWH
Abbotskerswell	30112312111233020211302000000010000000000000000000000000
Barton	22221211211114101012102000010001110000010000000000000000
Bovey Tracey	20122110122112120313202000010002100000000000000000000000

GROUND/UMPIRE INCIDENCE (full season)

	ABCDDHHJLMNPRRRSSVVWWBHEGGHHJMMRRSTCCFKLPRGHMPTWMHCQSSSWH
Abbotskerswell	10110201100221010101200000000001000000000000000000000000
Barton	02020201111112000011000000000000110000000000000000000000
Bovey Tracey	20101110111102100101101000010000000000000000000000000000

UMPIRE/UMPIRE INCIDENCE (full season)

	ABCDDHHJLMNPRRRSSVVWWBHEGGHHJMMRRSTCCFKLPRGHMPTWMHCQSSSWH
J Anning	01100110001110100101111000001000000000000000000000000010
S Bilverstone	10000101000101010000001100001000001001000100000000000000
G Curson	10000110000010011000001000000000000000100000001000000010

Each column refers to an umpire in the usual order.

Week-by-week usage of umpires

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
J Anning	P	P	A	P	B		A	P	P	X	A	P		P	A	P	P	P
S Bilverstone	B		A	A	P	A	P		B	P	A		P		P		A	
G Curson	X	P	P	X	A	P	X	P	B	A	P	X	X	X	X	P		P

Each column refers to a Saturday during the season. The Division is shown; X means that the umpire is not available on that date; a blank means available but not used.