European Journal of Operational Research 236 (2014) 592–600

Contents lists available at ScienceDirect

European Journal of Operational Research

journal homepage: www.elsevier.com/locate/ejor

Decision Support

Improved bounds for the traveling umpire problem: A stronger formulation and a relax-and-fix heuristic ☆

Lucas de Oliveira [a], Cid C. de Souza [a], Tallys Yunes [b,*]

[a] Institute of Computing, University of Campinas (UNICAMP), Campinas, SP, Brazil
[b] School of Business Administration, University of Miami, Coral Gables, FL, USA

ARTICLE INFO

Article history:
Received 29 August 2013
Accepted 11 December 2013
Available online 18 December 2013

Keywords:
OR in sports
Baseball
Heuristics
Integer programming
Relax-and-fix
Traveling umpire problem

ABSTRACT

Given a double round-robin tournament, the traveling umpire problem (TUP) consists of determining which games will be handled by each one of several umpire crews during the tournament. The objective is to minimize the total distance traveled by the umpires, while respecting constraints that include visiting every team at home, and not seeing a team or venue too often. We strengthen a known integer programming formulation for the TUP and use it to implement a relax-and-fix heuristic that improves the quality of 24 out of 25 best-known feasible solutions to instances in the TUP benchmark. We also improve all best-known lower bounds for those instances and, for the first time, provide lower bounds for instances with more than 16 teams.

1. Introduction

The assignment of officials (referees, umpires, judges, etc.) to the games of a competition is an important and difficult problem studied in the area of sports scheduling. The specific constraints and objectives vary according to the sport and type of competition, of course, but they typically aim to satisfy a given set of fairness criteria while minimizing costs (e.g. wages or travel). Existing research ranges over many different sports, including baseball (Evans, Hebert, & Deckro, 1984; Evans, 1988; Trick & Yildiz, 2011; Trick & Yildiz, 2012; Trick, Yildiz, & Yunes, 2012), cricket (Wright, 1991), football (Yavuz, İnan, & Fığlalı, 2008), and tennis (Farmer, Smith, & Miller, 2007). For more comprehensive surveys of sports-related problems, we refer to Ernst, Jiang, Krishnamoorthy, Owens, and Sier (2004) and Kendall, Knust, Ribeiro, and Urrutia (2010).

We study the traveling umpire problem (TUP), which was first proposed by Trick and Yildiz (2007) as an abstract version of the real-life umpire scheduling problem faced by Major League Baseball. Despite excluding many details present in the real problem, the TUP successfully captures the most important features that make the problem very challenging to solve. This is evidenced by the fact that many small instances remain unsolved in the official TUP benchmark: Trick (2013).

Given a double round-robin tournament with $2n$ teams (each team plays against each other team twice, once at home and once on the road, over exactly $4n-2$ rounds), the distances between the home venues of each pair of teams, and two integers $0 \leqslant d_1 < n$ and $0 \leqslant d_2 < \lfloor \frac{n}{2} \rfloor$, a solution to the TUP is an assignment of n umpire crews (umpires, for short) that satisfies the following constraints:

(i) In each round, each umpire is assigned to exactly one game and each game must be assigned to an umpire;
(ii) Each umpire visits every team at home at least once;
(iii) Each umpire visits any given venue at most once in any sequence of $n-d_1$ consecutive games;
(iv) Each umpire sees any given team at most once in any sequence of $\lfloor \frac{n}{2} \rfloor - d_2$ consecutive games.

The objective is to find a feasible solution that minimizes the total distance traveled by the umpires over the entire tournament.

When $d_1 = d_2 = 0$, TUP instances tend to be more difficult to solve because constraints (iii) and (iv) become stricter. We refer to these instances as *hard instances*. To allow for a wider range in the degree of difficulty, the TUP benchmark also includes instances with $d_1 + d_2 > 0$, to which we refer as *relaxed instances*.

☆ This research was supported by Grant 01-P-01965-2012 from *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (Brazil) and Grants 142278/2013-0, 477692/2012-5 and 302804/2010-2 from *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (Brazil).

* Corresponding author. Address: School of Business Administration, University of Miami, 5250 University Drive, Room KE405, Coral Gables, FL 33146, USA. Tel.: +1 305 284 5107; fax: +1 305 284 2321.

E-mail addresses: lucas.oliveira@ic.unicamp.br (L. de Oliveira), cid@ic.unicamp.br (C.C. de Souza), tallys@miami.edu (T. Yunes).

0377-2217/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ejor.2013.12.019

Our main contributions are: (1) we strengthen a known integer programming formulation for the TUP and use it to implement a relax-and-fix heuristic that improves the quality of 24 out of 25 best-known feasible solutions to instances in the TUP benchmark; and (2) using our stronger formulation, we improve all best-known lower bounds for those instances and, for the first time, provide lower bounds for instances with more than 16 teams.

Before explaining our approach in detail, we review the existing exact and heuristic methods for solving the TUP.

2. Previous work

Trick and Yildiz (2007) present exact integer programming (IP) and constraint programming (CP) models for the TUP and test them on benchmark instances ranging from 4 to 16 teams, as well as on a 30-team instance. These same models are also used in Trick and Yildiz (2011); Trick and Yildiz (2012); and Trick et al. (2012), but in Trick and Yildiz (2011) and Trick and Yildiz (2012) the performance of the IP model is improved by a better choice of solver parameters (execution times were limited to three hours). Their IP model finds optimal solutions to all instances with at most 10 teams. The CP model finds optimal solutions to all instances with at most 8 teams, and also to one of the four 10-team instances. In addition, it manages to prove that the 12-team instance is infeasible; a conclusion that was not obtained by the IP model within the allowed computation time. When it comes to hard instances with 14 teams, both IP and CP find feasible solutions to all four instances, with the CP model beating the IP model in terms of solution quality in three of the four cases. Neither the IP nor the CP models managed to find any feasible solutions to hard instances with more than 14 teams. When run on relaxed instances, the CP model finds feasible solutions to all eight 14-team and all twelve 16-team instances. The IP model finds feasible solutions to 17 of these 20 instances (three 16-team instances had no solution after 3 hours), but all 17 solutions are better than their counterpart solutions found by the CP model. Neither method managed to find provably optimal solutions to any of the relaxed instances with more than 10 teams.

On the heuristic side, we focus on hard and relaxed instances with at least 14 teams because none of them have known optimal solutions (29 instances in total; 25 with known feasible solutions). Trick et al. (2012) describe a greedy matching heuristic (GMH) and a two-exchange local search that are combined to build a simulated annealing (SA) heuristic. The SA heuristic does reasonably well on the real-life Major League Baseball problem of 2006, but it does not perform so well on the TUP instances. Trick and Yildiz (2011) incorporate the GMH into a large neighborhood search guided by Benders cuts that help repair the solution being built when the heuristic gets stuck. This improved GMH, which they call GBNS, finds feasible solutions to 23 out of 25 instances with previously known solutions, with 16 of those solutions being improvements over the best results at the time. In a follow-up paper, Trick and Yildiz (2012) propose a genetic algorithm (GA) with a crossover operator that uses a matching scheme to recombine the individuals of a population. Their GA further improves the quality of 14 instances with respect to the GBNS results. As of August, 2013, the results published in Trick (2013) indicate that the four best-known solutions to the 14-team hard instances were obtained by Wauters (2013). Table 1 shows how many of the 25 best-known solutions to date have been found by each of the most successful methods described above. The GA currently owns the majority of best-known solutions (13 out of 25), including the best solution to the 30-team instance.

The remainder of this paper is organized as follows. In Section 3, we present the IP formulation of Trick and Yildiz (2011) and show how it can be strengthened. We describe our relax-and-fix heuristic in Section 4, and provide computational results in Section 5. We conclude the paper and discuss future research directions in Section 6.

3. IP formulations

For simplicity, we use letters i and j to refer to teams i and j, as well as their respective home venues. In addition, we use letters u and s to refer to an umpire and a round in the tournament, respectively.

3.1. The original formulation

The IP formulation used by Trick and Yildiz (2011); Trick and Yildiz (2012); and Trick et al. (2012) starts with the following input data:

- Set of umpires $U = \{1, \ldots, n\}$;
- Set of teams $T = \{1, \ldots, 2n\}$;
- Set of rounds $S = \{1, \ldots, 4n-2\}$;
- $\mathrm{OPP}[s,i] = \begin{cases} j & \text{if } i \text{ plays against } j \text{ at venue } i \text{ in round } s \\ -j & \text{if } i \text{ plays against } j \text{ at venue } j \text{ in round } s; \end{cases}$
- d_{ij} = distance in miles between venues i and j;
- $CV_s = \{s, \ldots, s+n-d_1-1\}$ for any given round $s \in \{1, \ldots, 4n-2-(n-d_1-1)\}$;
- $CT_s = \{s, \ldots, s+\lfloor \frac{n}{2} \rfloor - d_2 - 1\}$ for any given round $s \in \{1, \ldots, 4n-2-(\lfloor \frac{n}{2} \rfloor - d_2 - 1)\}$.

The decision variables are:

- $x_{isu} = \begin{cases} 1 & \text{if the game at venue } i \text{ inround } s \text{ isassigned} \\ & \text{toumpire } u \\ 0 & \text{otherwise}; \end{cases}$
- $z_{ijsu} = \{ 1$ if umpire u is at venue i in round s and travels to venue j in round $s+1 0$ otherwise.

We are now ready to state the formulation.

$$\min \sum_{i\in T}\sum_{j\in T}\sum_{u\in U}\sum_{s\in S:s<|S|} d_{ij} z_{ijsu} \tag{1}$$

$$\sum_{u\in U} x_{isu} = 1, \quad \forall\, i \in T, s \in S : \mathrm{OPP}[s,i] > 0, \tag{2}$$

$$\sum_{i\in T:\mathrm{OPP}[s,i]>0} x_{isu} = 1, \quad \forall\, s \in S, u \in U, \tag{3}$$

$$\sum_{s\in S:\mathrm{OPP}[s,i]>0} x_{isu} \geqslant 1, \quad \forall\, i \in T, u \in U, \tag{4}$$

$$\sum_{c\in CV_s:\mathrm{OPP}[c,i]>0} x_{icu} \leqslant 1, \quad \begin{array}{l} \forall\, i \in T, u \in U, s \in S : \\ s \leqslant |S| - (n-d_1-1), \end{array} \tag{5}$$

$$\sum_{c\in CT_s} \left(x_{icu} + \sum_{j\in T:\mathrm{OPP}[c,j]=i} x_{jcu} \right) \leqslant 1, \quad \begin{array}{l} \forall\, i \in T, u \in U, s \in S : \\ s \leqslant |S| - (\lfloor \frac{n}{2} \rfloor - d_2 - 1), \end{array} \tag{6}$$

$$x_{isu} + x_{j(s+1)u} - z_{ijsu} \leqslant 1, \quad \forall\, i,j \in T, u \in U, s \in S : s < |S|, \tag{7}$$

$$x_{isu} \in \{0,1\}, \quad \forall\, i \in T, u \in U, s \in S, \tag{8}$$

$$z_{ijsu} \in \{0,1\}, \quad \forall\, i,j \in T, u \in U, s \in S : s < |S|. \tag{9}$$

The objective function (1) minimizes the total distance traveled by the umpires. Constraints (2) and (3) state that each game is refereed by an umpire, and each umpire is assigned to a game, respectively. TUP constraints (ii), (iii), and (iv) from Section 1 are modeled by (4)–(6), respectively. Finally, (7) ensures that game (x) and travel (z) assignments are consistent.

Trick and Yildiz (2011) improve the above formulation by including the following constraints: