Grouping and timetabling for multi-league sports competitions

Miao Li[1][0000−0002−6205−814X] and Dries Goossens[1][0000−0003−0224−3412]

Faculty of Economics and Business Administration, Ghent University,
Tweekerkenstraat 2, 9000 Ghent, Belgium
miao.li@ugent.be;dries.goossens@ugent.be

Abstract. This work presents a bi-objective grouping and timetabling problem for sports competitions that are played using multiple leagues. We propose a decision-making framework to uncover the trade-off between minimizing travel distance and venue capacity violations, in order to meet the preferences of different stakeholders.

Keywords: OR in sports · bi-objective optimization · sport team grouping · sports timetabling · multi-league scheduling

1 Introduction and problem description

Youth and amateur sports provide non-professional players with the opportunity to exercise and develop athletic skills. In such sports, the number of teams involved can reach the thousands, and considerable efforts are required to organize their competition. A first challenge is that teams with players of the same age, gender or strength need to be grouped into leagues. This problem is known as the sports team grouping problem (STGP [3]). The main objective with this problem is to minimize the total travel distance travelled by all teams, knowing that teams visit each other team in their league, but none of the teams from other leagues. Another issue is setting up a timetable for each of these leagues, i.e. deciding when each match is to be played. This problem is known as multi-league sports timetabling problem (MLSP [1]). Since teams of the same club share the same infrastructure (venue), whose capacity should be respected, the leagues are interdependent and cannot (optimally) be scheduled one by one. In particular, per time slot on which the number of home-playing teams from a club exceeds the number of home matches the club can host, a capacity violation arises. Hence, the main objective in the multi-league sports timetabling problem is to minimize the total venue capacity violations over all clubs.

In practice, these problems are handled sequentially: first solve the sports team grouping problem, and then, based on the resulting league composition, solve the multi-league sports timetabling problem. However, some team grouping may allow a timetable where few or even no clubs face a capacity issue, while another grouping may be more problematic, resulting more venue capacity violations. As its main novelty, this paper integrates both decision problems

(i.e., STGP and MLSP) as the multi-league grouping and timetabling problem (MLGTP), with two objective functions: the minimization of total distance travelled by all teams and the minimization of total venue capacity violations over all clubs. Both objectives are demonstrated to be conflicting when leagues have different sizes. We therefore investigate their trade-off, and develop a method that allows us to approximate the Pareto front.

2 Proposed method

For instances with a somewhat realistic scale, an approach like the ϵ-constraint method (ECM) [2] is intractable within a reasonable computation time. Hence we develop a two-phase two-layer constructive algorithm (SLCM) to find an approximate Pareto front for our bi-objective problem in a reasonable time. The optimization process begins with an initial Pareto solution. Next, in the first phase, the problem is decomposed into STGP and MLSP sub-problems, which are solved by a two-layer method sequentially and iteratively. At each iteration, the outer layer is first used to minimize the total distance travelled, where the initial assignment of teams to leagues is further improved by simulated annealing. In the second phase, we enlarge the search space and improve incumbent candidate Pareto solutions. Then, given a list of potential Pareto efficient solutions of MLGTP, an approximate Pareto front is identified. The overall process is able to optimize two distinct objectives simultaneously.

3 Preliminary results and conclusion

Due to the fact that there are no MCGTP instances available in the literature, we created some sets of instances with various numbers of teams, leagues and clubs, as well as league and club sizes, as given in Table 1.

Table 1. Overview of instance types and their features

Instance type	Instance ID	No. teams	No. clubs	No. leagues	League size				
					16	10	8	6	4
Small-scale	S1	18	8	3	–	–	1	1	1
	S2	34	16	3	1	1	–	1	–
Large-scale	N1	80	17	8	2	–	6	–	–
	N2	112	18	11	3	–	8	–	–
	N3	144	20	13	5	–	8	–	–
	N4	176	20	16	6	–	10	–	–
	N5	208	25	19	7	–	12	–	–

Note: '–' indicates the instance type does not contain leagues of the corresponding size

After preliminary tests for parameter configurations, we evaluate the performance of the proposed bi-objective solution method. Besides the two-phase two-layer constructive method (SLCM), in order to study the contribution of the second phase, each problem instance is solved by the two-layer constructive method (LCM) which only includes the first phase of SLCM.

We first assess the ability of the methods on small-scale instances. The results show that SLCM is capable of producing optimal Pareto solutions for all instances, and that the approximate Pareto set obtained by the LCM lies close to the optimal set. Additionally, compared with the computationally demanding ECM, LCM and SLCM both operate far more efficiently.

We next turn our attention to large-scale instances. For these instances, ECM cannot obtain a feasible solution within the time limit (7200s). Four evaluation metrics, namely the number of Pareto efficient solutions (NPS), the diversification metric (DM), the mean ideal distance (MID) and the spacing metric (SM) are applied to compare different sets of Pareto solutions produced by SLCM and LCM. While with respect to DM, both algorithms display more or less the same performance, based on the metrics NPS, MID and SM, SLCM is clearly superior to LCM. Overall, this indicates that the second phase of the SLCM approach is a valuable addition.

With respect to computation time, small-scale instances were solved by SLCM in less than 405 seconds, while we needed no more than 1800 seconds for the large-scale instances.

In summary, our preliminary results suggest that the proposed method is able to offer a good approximation of the Pareto front, helping the league organizers to find a good compromise proposal to balance the travel distance of all teams and the venue capacity violations over all clubs.

References

1. Davari, M., Goossens, D., Beliën, J., Lambers, R., Spieksma, F.C.: The multi-league sports scheduling problem, or how to schedule thousands of matches. Operations Research Letters **48**(2), 180–187 (2020)
2. Haimes, Y.: On a bicriterion formulation of the problems of integrated system identification and system optimization. IEEE transactions on systems, man, and cybernetics **1**(3), 296–297 (1971)
3. Toffolo, T.A., Christiaens, J., Spieksma, F.C., Vanden Berghe, G.: The sport teams grouping problem. Annals of Operations Research **275**(1), 223–243 (2019)