MILP Based Approaches for Scheduling Double Round-Robin Tournaments

Daniil Sumin · Ivan Rodin

Abstract This paper presents some mixed-integer linear programming (MILP) based models for solving the International Timetabling Competition on Sports Timetabling (ITC2021). The approach explained here came third in the competition, and it found the best solution in 16 out of the 45 instances.

Keywords International Timetabling Competition 2021 · Mixed-Integer Linear Programming · Sports Timetabling · Home-Away Patterns · Decomposition

1 Introduction

This paper presents the approaches used to solve the ITC2021 timetabling problems [2]. The article is organized as follows. Problem description is given in Section 2. Section 3 introduces the Baseline model that is MILP monolithic formulation for scheduling problem. Section 4 describes the Patterns model consisting of two optimization parts: possible patterns generation and assigning patterns to the teams. Section 5 proposes the Patterns Mirrored model that is based on the mirroring format of the some real-life competitions. Section 6 presents the most applicable model called the 2-Phased model that decomposes the problem into two consecutive parts: the first and the second round of the competition. A combination of the Patterns and the 2-Phased approaches is described in Section 7. Proposed models' applicability for different types of instances is considered in Section 8.

D. Sumin
Higher School of Economics, Russia, Moscow
E-mail: danilsum21@gmail.com

I. Rodin
Moscow State University, Russia, Moscow
E-mail: del.ris@yandex.ru

2 Problem description

The input data is a set of settings and constraints of optimization problems specified in the RobinX format [4] (although not all features and constraints of this format were included in ITC2021 instances). All constraints are divided into two types: hard and soft. Hard constraints can never be violated, while maximum number of the soft constraints should be satisfied with respect to their importance. The main goal of the work is to develop an algorithm for scheduling a double round-robin tournament in which all hard constraints are satisfied and the penalty for soft constraints is minimized.

On the other hand, constraints can be divided into the following groups: capacity constraints (CA1,CA2,CA3,CA4), game constraints (GA1), break constraints (BR1,BR2), fairness constraints (FA2) and separation constraints (SE1). Capacity constraints regulate when teams play home or away in a specific group of slots. Game constraints are used to forbid or to force playing specific games in certain slots. Break is a situation when a team plays two consecutive games with the same home or away status. Break constraints are used to manage the quantity of such situations. Nurmi et al. in [3] proposed a measure of fairness of a sports competition called "k-balancedness" which requires the difference in played home and away games to be smaller than k at any point of the season. There is no sign of FA2 constraints in all proposed models, because FA2 requires a lot of additional variables and constraints. FA2 constraints are implicitly optimized by BR2 and hard CA3. As a result, there are almost no violations of soft FA2 constraints in all instances. Separation constraints are used to control the interval between two games with the same opponents.

3 Baseline model

Baseline model is a full MILP formulation of the ITC2021 problem. This model considers a binary decision variable z_{s,t_1,t_2} that is equal to 1 if team t_1 plays at home against team t_2 on time slot s and 0 otherwise. The model is based on [4], where authors mathematically formulated all constraints following the RobinX data format. Some constraints in [4] were nonlinear, therefore, standard linearization techniques were used in the Baseline model.

This approach is developed for instances with the following properties:

1. small-sized problems;
2. problems without hard BR2 constraints;
3. problems without any type of SE1 constraints.

4 Patterns model

Breaks are one of the bottlenecks of the Baseline model, because they require a lot of linearization constraints and decision variables, so we propose the

Patterns model. The sequence of home and away games is called team's home-away pattern (HAP). The Patterns model consists of two optimization stages.

At the first stage we minimize the number of breaks in a competition using $h_{s,\,t}$ binary decision variables which indicate that team t plays at home on time slot s. Some constraints can be taken into account at this stage: total number of home games for each team, total number of home teams on each time slot, hard CA1 constraints, hard CA3 constraints and a subset of hard GA1 constraints. The set of possible patterns contains the solution of the first stage and all other patterns with zero or one break.

The second stage is similar to the Baseline model, apart from dealing with breaks. It has additional binary decision variables $q_{t,\,p}$ which indicate that team t follows HAP p from the set of possible patterns.

This approach is developed for instances with the following properties:

1. small-sized problems;
2. problems without any type of SE1 constraints;
3. problems where the biggest contribution to the objective function is made by soft BR2 constraints.

5 Patterns Mirrored model

Some European football leagues use a mirrored competition format, where the second half of the competition is identical to the first one with an inverted home advantage. The Patterns Mirrored model uses both stages of the Patterns model along with mirroring constraints. Mirrored format allows us to reduce the problem size significantly, because the schedule of the second round is completely defined by the schedule of the first. But with the mirrored scheme, we considerably reduce the feasible region, so some instances can be infeasible with this approach. Mirrored format automatically satisfied both separation and phased constraints. Moreover, we should take into account the lower bound on the number of breaks in the mirrored double round-robin tournament $3T-6$ [1], where T is the number of teams in competition.

This approach is developed for instances with the following properties:

1. large-sized problems;
2. problems with any type of SE1 constraints.

6 2-Phased model

2-Phased model is the decomposition approach for the ITC2021 problem. The idea is to divide the solution into two consecutive stages. At the first stage we build the schedule of the first phase of competition. At the second stage we schedule the second phase with respect to the solution of the first stage. Some of the constraints can be independently divided into two rounds: CA3, BR2 and SE1. In Phase 1, we tighten mutual constraints (CA1, CA2, CA4, GA1, BR1) for having an opportunity to satisfy them while solving Phase 2. For these constraints, *max* constant from the respective constraint is distributed proportionally between two phases with respect to the number of slots of a certain phase in constraint. Example for the competition with 18 teams:

Initial constraint	CA2 (max = 2) with $slots = \{0; 1; 20; 21\}$
Generated the first phase constraint	CA2 (max = 1) with $slots = \{0; 1\}$
Generated the second phase constraint	CA2 (max = 1) with $slots = \{20; 21\}$

Besides, two special constraints are included in the first stage: if we have a mandatory game between team 1 and team 2 in the second round of the competition from GA1, then we know about their home-away status in the first round and the set of possible first round slots for their game for satisfying the SE1 constraint.

This approach is developed for instances with the following properties:

1. large-sized problems;
2. problems with any type of SE1 constraints;
3. problems, where the Patterns Mirrored model is infeasible.

7 Patterns 2-Phased model

The 2-Phased model could have trouble with satisfying the hard BR2 constraints, and may also incur high penalties in the objective function due to the soft BR2 constraints. Combining the Patterns model and the 2-Phased model into Patterns 2-Phased model can improve the results. The combined model consists of all the features of the two models: generation of possible patterns, and application of two consecutive phase models for assigning patterns to the teams.

This approach is developed for instances with the following properties:

1. large-sized problems;
2. problems with any type of SE1 constraints;
3. problems with strong influence of break constraints.

8 Applicability

We can say that the model applicability area is a class of problems, where this approach can find feasible solution in a reasonable time. Each model was developed to handle certain types of instances, but the scope of the approach can be much larger. Models applicability is shown in Fig. 1. The Baseline model is not applicable to large problems, instances with hard BR2 constraints, and instances with SE1 constraints. Patterns model can be applied to problems with hard BR2 constraints. Other models can be applied to any type of problem.

The general algorithm for solving each of the instances in the competition can be described as follows:

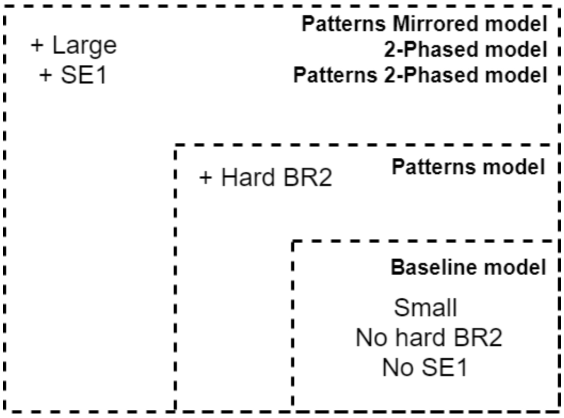

Fig. 1 Models' applicability

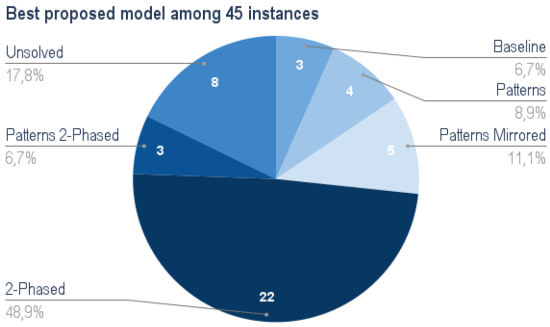

Fig. 2 Best proposed model among 45 instances

1. Choose appropriate formulations to apply to the problem instance to be solved depending on which constraints are present in the instance.
2. Solve the problem instance using the formulations selected in Step 1.
3. Select the best solution from the set of all produced solutions.

All instances from ITC2021 were solved according to the algorithm described above. Fig. 2 shows the distribution of the best solutions provided by the different formulations. It can be seen that the 2-Phased model is the most applicable approach for solving ITC2021 instances. In 7 out of 8 unsolved instances, the hard BR2 constraints could not be satisfied that is motivation for future research.

References

1. De Werra, D. (1981). Scheduling in sports. Studies on graphs and discrete programming, 11, 381-395.

2. Van Bulck, D., Goossens, D., Beliën, J., and Davari, M. (2021). The Fifth International Timetabling Competition (ITC 2021): Sports Timetabling. Proceedings of MathSport International 2021 Conference, MathSport, pp. 117-122.
3. Nurmi K, Goossens D, Bartsch T, Bonomo F, Briskorn D, Duran G, et al. A framework for scheduling professional sports leagues. In: Ao S, Chan A, Katagiri H, Xu L, editors. AIP Conference Proceedings. AMER INST PHYSICS; 2010. p. 14–28.
4. Van Bulck D, Goossens D, Schönberger J, Guajardo M. RobinX : a three-field classification and unified data format for round-robin sports timetabling. EUROPEAN JOURNAL OF OPERATIONAL RESEARCH. 2020;280(2):568–80.