Rescheduling the NBA regular season via Integer Programming

Juan José Miranda Bront[1,2] and Nicolás García Aramouni[1]

[1]Universidad Torcuato Di Tella, Buenos Aires, Argentina
[2]Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina

Abstract

The COVID-19 pandemic generated disruptions across multiple sectors, in particular in the sports industry as many leagues had to re-adapt their competitions as a consequence of lockdowns, travel restrictions and other implemented safety measures. Even when activities started to resume, match suspensions continued throughout the different leagues. Dealing effectively with these unexpected events is extremely relevant regarding both sports and economic aspects of the competitions. In this paper, we propose a framework that builds upon Integer Programming models to systematically reschedule suspended games and generate a contingency fixture that accounts for relevant operational constraints. Using the 2020-21 NBA season as a benchmark, we compare different scheduling approaches under two different objective functions. Computational results show that the approach produces good quality schedules and that the framework has potential to be applied in practice.
Keywords: sports scheduling; integer programming; tournament rescheduling

1 Introduction

The COVID-19 pandemic has been one of the most challenging events the humanity faced in the last decades. This unexpected event stressed the nationals' health systems across the world, generating millions of cases and deaths, as well as the world economy, that was affected the health and safety protocols implemented. Estimations indicate that global GDP decreased 3.3% in 2020 [1] (as a point of comparison, this number was 1.3% during the global crisis of 2009), and that global unemployment rose from 5.4% to 6.6% in 2020 [2]. The sports industry was not the exception, as most major tournaments undergone an initial indefinite interruption during March-April 2020. Later on, some competitions started to resume their activities with new formats, and generally, without fans on the stadiums. For instance, the uncertainty introduced by the pandemic motivated some tournaments to be played in a "bubble" format to finish the 2019-20 season. where the teams played all games in a central location (including mobility restrictions for players as well) to prevent the teams from traveling among different locations. This is the case of multiple leagues across different sports, such as football (UCL), basketball (NBA) and ice hockey (NHL).

Despite the lack of attendance to the stadiums, the sports leagues faced important challenges throughout this period. Players, coaches and the management could still get sick and sidelined due to the safety protocols. Depending on the number of players affected, some matches could get suspended.

This situation generated new challenges for league organizers, as a large number of suspensions would affect the execution of the schedule. In pre-pandemic contexts, suspensions were rare and unrelated to each other (generally related to a specific reason). On the contrary, during the COVID19 pandemic infected people required several days to recover and clear the safety protocols. In this scenario, it is likely that multiple consecutive matches might get suspended. This is specially problematic for time-relaxed schedules, where generally multiple matches per week are played. For the 2020-21 NBA season, the probability of a team's game being rescheduled given that the previous one was rescheduled was 33.8%. To reduce the impact, different actions were considered. For instance, the NBA allowed teams to sign 10-day contracts with free agents, that acted as temporary replacements for players entering the protocols. If necessary, matches were also suspended and rescheduled to be played later in the season. These contingency plans affect the planning, but also impact the competition.

To our knowledge, the literature on reschedudling strategies in sports competitions is rather scarce. Yi et. al. [3] present both proactive and reactive strategies to reschedule football matches in a pre-pandemic scenario for time-constrained tournaments. However, we are not aware of research involving rescheduling in a time-relaxed setup. Based on this preliminary analysis, in this paper we explore different strategies to systematically adapt time-relaxed schedules when affected by games suspensions due to unexpected or uncontrolled events during its execution. More specifically, we propose two alternative strategies, that build on integer linear programming models to reschedule suspended matches within the original, planned schedule respecting different rules, such as the number of consecutive games, the distance travelled, among others. We consider two different potential objective functions that might be considered by league organizers, and compare the alternative rescheduling strategies using different quality metrics through computational experiments. We concentrate on the 2020-21 NBA season to explore our models, although our approach could fit in other contexts.

2 Problem definition

In this section we provide a high level description of the key rules for our approach. We assume the tournament has an *original schedule* for the season, indicating the planned date and location for each game throughout the planning horizon. During the execution, due to either unexpected or uncontrollable events, some games need to be suspended from their scheduled date and postponed to be played later in the season. We further assume that all suspended games need to be played to conclude the season, eventually after the final date in the original schedule. Indeed, a somehow straightforward alternative is to wait until the end of the season and reschedule all the suspended games afterwards, assuming there is enough time available before the next stages of the competition (e.g., playoffs). Such an approach, however, may affect some specific teams due to fairness issues, specially considering that their final standings would by affected by the fact that a large number of rescheduled games should be played in a short period of time and, potentially, right before decisive stages of the competition.

We formulate the problem in a more general fashion. Assume the schedule has been executed up to some specific time in the original planning horizon (eventually, the end of the season), and that some games have been suspended. Intuitively, the (sports) timetable rescheduling problem involves finding feasible dates for the cancelled games in the rest of the season. In this work, we aim to insert the suspended games into the schedule without modifying the game dates for matches from the original schedule.

Following the definitions introduced in [3] for time-constrained tournaments, also present in other rescheduling contexts, each game in the executed timetable that is played before or after its scheduled round is considered a *disruption*. In our setup, each suspended game will therefore translate into a disruption that needs to be rescheduled in the remaining of the season schedule. Then, the challenge is to find feasible rounds to reschedule each disruption, obtaining a feasible rescheduled timetable.

Given a disrupted game, a date t (round) is considered a candidate if the following conditions hold:

1. t occurs after the disruption's date in the original schedule;
2. there are no planned games for the teams involved in the disruption on date t;
3. there are no scheduling rules violations if the corresponding match is scheduled on t; and
4. both teams travel a *reasonable* distance if the corresponding match is scheduled on t.

Even in a time-relaxed system, there are several rules that are relevant in order to create a reasonable and fair schedule. Condition 3 goes in that direction. For example, we impose that no team should have games on three consecutive days to make the schedule realistic regarding rest times between games. Therefore, if a team has matches on Match 13^{th} and 14^{th}, no disrupted game should be rescheduled on March 15^{th}. We carry out this analysis, evaluating the number of matches (both home, away and in general) that are being held on a set of consecutive 1 to 7 days. Another relevant aspect involves considering the total distance travelled, one of the most common metrics used to evaluate sports schedules. For example, if a match held in Los Angeles has to be rescheduled (let's say between the Lakers and the Rockets), it could make sense to avoid using a date during a tour in the east coast, for instance after a game in New York and before a game in Boston. Thus, we prioritize dates that generate traveled distances closer to the ones incurred in the case the match was not suspended. Condition 4 incorporates this aspect to obtain a new fixture having reasonable tours.

We note that there are situation where some disruptions may not be rescheduled within the rest of the season even when these sets contain feasible dates. In that case, we determine that match would be played after the last game of each team, respecting that no franchise should play on three consecutive days, and potentially adding new dates to the competition if needed, in order to play these games.

3 Mathematical Model

We first introduce some basic definitions and notation used to model the problem. Let $S = \{1, 2, \ldots, m\}$ be the set of teams (in our case, NBA teams), and $T = \{1, 2, \ldots, r\}$ the set of original tournament days (rounds). Let (j, k) denote a match between teams $j, k \in S$, and we further represent a *scheduled game* by a pair $\alpha =< (j, k), t >$ indicating that game (j, k) takes place on day $t \in T$. With these definitions, the schedule for the season can be modelled as a set R of games.

Given a team $i \in S$, we define $R_i, R_i^{\text{dis}} \subseteq R$ as the set of scheduled games (i.e., do not need to be rescheduled) and the set of original disrupted matches for team i (i.e., that need to be rescheduled), respectively. Let $R^{\text{dis}} = \cup_{i\in S} R_i^{\text{dis}}$ denote the set of all the disrupted matches, including their original date, to be rescheduled. For every disruption game α, let T_α^{free} as the set of potential candidate dates in Ts to reschedule match α. In order to incorporate the conditions imposed to the schedule, given $t_1, t_2 \in T$, $t_1 < t_2$, let MG_{t_1,t_2} indicate the maximum number of games a team can play within every window of $t_2 - t_1$ days. In our case, we consider sliding windows of $1 \leq t_2 - t_1 \leq 7$ days, each of them with value corresponding value for MG_{t_1,t_2}. Finally, let $k^i_{t_1,t_2}$ denote the number of games originally scheduled for team $i \in S$ between t_1 and t_2. For each disruption $\alpha \in R^{\text{dis}}$ and $t \in T_\alpha^{\text{free}}$ we define binary variables $x_{\alpha t}$ that are equal to 1 if the new date for match α is t. The ILP mathematical model reads:

$$\max \sum_{\alpha \in R^{\text{dis}}} \sum_{t \in T^{\text{free}}_{\alpha}} x_{\alpha t} \tag{1}$$

$$\text{s.t.} \sum_{t \in T^{\text{free}}_{\alpha}} x_{\alpha t} \leq 1 \qquad \forall \alpha \in R^{\text{dis}} \tag{2}$$

$$\sum_{\alpha \in R^{\text{dis}}_i} \sum_{\substack{t_1 \leq t \leq t_2, \\ t \in T^{\text{free}}_{\alpha}}} x_{\alpha t} + k^i_{t_1,t_2} \leq MG_{d_{t_1,t_2}} \qquad \forall t_1, t_2 \in T, 1 \leq t_2 - t_1 \leq 7, i \in S \tag{3}$$

$$x_{\alpha t} \in \{0, 1\} \qquad i \in R^{\text{dis}}, t \in T^{\text{free}}_{\alpha} \tag{4}$$

The objective function (1) maximizes the number of matches rescheduled within the original set of dates of the season. We refer to this objective function as MaxG. Constraints (2) force each match to be rescheduled at most once during the original schedule dates. Constraints (3) enforce the new schedule satisfies Condition 3, while constraint (4) defines the variables to be binary.

In addition, for $\alpha \in R^{\text{dis}}$ and $t \in T^{\text{free}}_{\alpha}$, we define $d_{\alpha t}$ that indicates the number of days between the date in the original schedule for α and the potential new date, t. For experimental purposes, we consider the additional objective that minimizes the sum day difference between the original date and the new date. which reads

$$\min \sum_{\alpha \in R^{\text{dis}}} \sum_{t \in T^{\text{free}}_{\alpha}} d_{\alpha t} x_{\alpha t} \tag{5}$$

We refer to objective function (5) as MinD. We remark that constraints (2) are set as equality, rescheduling every disrupted match. In this case, only matches $\alpha \in R^{\text{dis}}$ with $T^{\text{dis}_{\alpha}} \neq \emptyset$ are considered.

It is important to mention that these formulations differ from other ILP models, such as the traveling tournament problem (TTP). Since our objective is to modify the original schedule as little as possible, the only we only consider decision variables to reschedule suspended matches. Thus, the constraints aim to impose some desired quality conditions on the new generated, even when some feasible dates may exist for some disrupted games. In all cases, the matches that cannot be rescheduled before the end of the season are scheduled afterwards, but still satisfying conditions 1 - 3.

4 Preliminary experimental results

We conducted computational experiments to analyze the differences among the planned NBA 2020-21 season schedule, the executed NBA 2020-21 season schedule, and scheduled obtained following the ideas presented in sections 2 and 3. During this season, 31 games were rescheduled due to COVID19 protocols. The mathematical models are implemented using Python 3 and CPLEX as an ILP solver.

We consider the following strategies implementing the model. For each strategy, we consider the MaxG and the MinD models.

- NBA exec: the timetable executed by the NBA. Recall that the schedule for the second half (after the All Star game) was defined from scratch including the suspended games form the first half. Thus, in addition to the benchmark with the implemented schedule, it provides a comparison with a proactive strategy with almost all the information known in advance.

- *monthly*: games are rescheduled on a monthly basis. We generate batches with the games suspended in a given month, which are rescheduled at the beginning of the next one throughout the remainder of the season. For the upcoming months, this new schedule is used as input.
- *monthly**: the *monthly* strategy with no distance restrictions. Also, if a game was rescheduled by the NBA within the month, we use that date.
- *Post All-Star*: we generate a unique batch including all the suspended games between the beginning of the season and the start of the All-Star Weekend, which are rescheduled in the remainder of the season. Suspended games after the All Star are rescheduled using the *monthly* strategy.
- *Post All-Star**: the *Post All-Star* adapted accordingly to *monthly**.

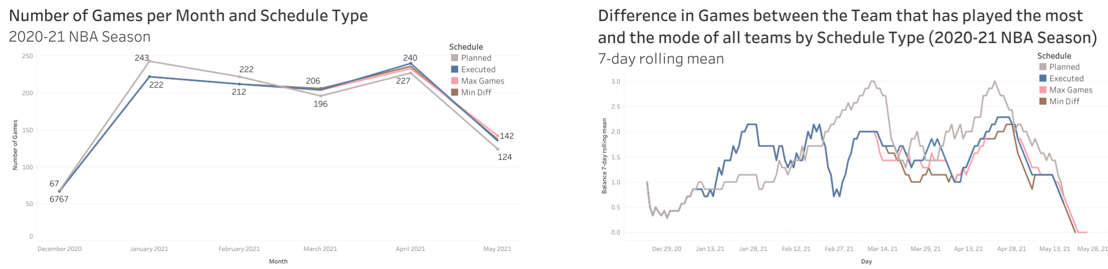
We consider first the instance provided by the scenario faced by the NBA during the 2020 - 2021 regular season, using as inputs the schedule and those games that were indeed suspended due to COVID protocols. Each combination between model and strategy, including the executed NBA season, is evaluated considering the total distance travelled and the number of breaks. For our approaches, since they are reactive strategies, we further report the number of additional rounds and games scheduled after the end of the season according to NBA exec. In Table 1 we report the percentage difference of MaxG and MinD with respect to the planned NBA schedule (i.e., with no disruptions).

metric / method	NBA exec	*monthly*		*monthly**		*Post All-Star*		*Post All-Star**	
		MaxG	MinD	MaxG	MinD	MaxG	MinD	MaxG	MinD
distance	-0.2%	0.9%	1.0%	1.5%	0.8%	1.3%	0.8%	1.3%	0.4%
breaks	0.6%	-0.2%	-0.1%	-0.3%	0.2%	-0.5%	0.2%	-0.4%	0.4%
# dates added	-	11	7	8	4	8	5	6	3
games after	-	14	8	9	3	15	8	10	3

Table 1: Comparison for the different strategies, relative to the NBA planned schedule.

The main message in Table 1 is that there are no major differences over these metrics when compared to the planned NBA schedule. Then, our approach is competitive compared to the approach used by the NBA, with all metrics fluctuating around a difference of 1%. One potential reason for this behavior may be the few potential reschedule options for each suspended match, considering that teams have, on average, two days between consecutive games. We note that, in all cases, MinD reschedules less games after the end of the season than MaxG, requiring also less additional rounds. In addition, we observe in almost all scenarios an increase on the distance travelled and an improvement in the total breaks. Finally, the comparison between each strategy with its relaxed version (i.e., *monthly* vs. *monthly** and *Post All-Star* vs. *Post All-Star**) shows that Condition 4 has a significant impact regarding the number of additional rounds needed.

In Figure 1, we also evaluate the distribution of games per month and the balance level that is generated by each schedule to obtain deeper insights on these results. Figure 1a shows the number of games per months for the NBA original schedule, the NBA executed schedule, and for the *monthly** strategy under the MinD and the MaxG objectives. As expected, our solution yields similar results than the ones that were presented by the NBA, as basically the same number of games are played in all scenarios. However, it is interesting to observe that the NBA originally planned for a more concentrated schedule on the first half of the season, leaving some flexibility for second half to eventually reschedule suspended games by implicitly using these *buffers* defined at the planning stage. This proactive



(a) Number of games per month by schedule.

(b) Balance of the schedule, *monthly** strategy.

Figure 1: Number of matches by month and balance level by schedule type - *monthly**

action is consistent with some recommendations obtained by [3] for time-constrained schedules. It is important to mention that this worked on practice because the amount of suspensions due to COVID cases peaked during December - February and decreased afterwards.

Figure 1b analyzes the *balance* of the different schedules, computed as a 7-day rolling mean between the team that played the greater number of games up to that point and the mode of all teams, for the *monthly** strategy. This visualization shows the differences among the schedules depending on the objective function used. We highlight the high variability in the original NBA planned schedule.

These results are heavily influenced by the particular disruptions that happened during the 2020-21 NBA season. More generally, it is difficult to assess whether these strategies, including the NBA executed schedule, would have worked in a context where the COVID19 cases remain constant over time. Thus, we generated larger instances, using the real season as input, we consider larger instances with more disruptions created as follows: (i) 15 additional disruptions, randomly selected throughout the tournament (50% + disruptions); (ii) 25 additional disruptions, randomly selected throughout the tournament (80% + disruptions); and (iii) 15 additional disruptions, all selected from games in March, modeling a second COVID19 wave.

instance / metric	distance	breaks	# dates added	games after
15 more games	3.3%	-0.3%	9	15
25 more games	5.8%	-0.2%	9	20
15 more games in March	1.8%	0.5%	9	17

Table 2: Difference against the planned NBA schedule, *monthly** strategy and MinD objective.

Table 2 shows the results for the *monthly** strategy using the MinD objective for these three additional scenarios, relative to the executed NBA schedule. Briefly, this experiment suggests that under more stressed scenarios, the impact on the different metrics may not be negligible, affecting eventually the business and the fairness of the competition.

References

[1] Gdp growth (annual %). *World Bank.*, 2022.

[2] Unemployment, total (% of total labor force) (modeled ilo estimate). *World Bank.*, 2022.

[3] X. Yi, D. Goossens, and F. T. Nobibon. Proactive and reactive strategies for football league timetabling. *European Journal of Operational Research*, 282(2):772–785, 2020.