

Maximizing the Attractiveness of a Football Tournament via Integer Programming

Gustavo Angulo¹, Alejandro Cataldo¹, Fernando Duarte¹, Nicolás Sanhueza¹,
Pablo A. Rey², and Antoine Sauré³

¹ Escuela de Ingeniería, Pontificia Universidad Católica de Chile, Santiago, Chile

² Departamento de Industria, Universidad Tecnológica Metropolitana, Santiago, Chile

³ Telfer School of Management, University of Ottawa, Ottawa, Canadá

Resumen We present a novel integer programming approach for scheduling football tournaments to maximize attractiveness. Our objective function keeps competition open until the final rounds by evaluating, at each matchday, the best and worst achievable standings for every team. We define five metrics to assess tournament attractiveness ex post, capturing aspects such as the number of teams that can still reach top or bottom positions and the uncertainty in overall standings. Applied to 24 real-league instances, the model yields significant improvements over historical schedules. Computational challenges arise from the rapid growth of the problem with the number of teams, pointing to the need for decomposition algorithms. We conclude that dynamically ordering matches can capture substantial additional attractiveness.

Keywords: sport scheduling , tournament attractiveness , integer programming , football , dynamic scheduling

Maximizando el atractivo de un torneo de fútbol mediante programación entera

Resumen Presentamos un nuevo enfoque basado en programación entera para la calendarización de torneos de fútbol con el objetivo de maximizar su atraktividad. Desarrollamos una función objetivo que evalúa, en cada fecha, los mejores y peores escenarios de clasificación posibles para cada equipo, manteniendo la competencia abierta hasta las últimas jornadas. Definimos además cinco métricas para evaluar la atraktividad de cualquier calendario, capturando aspectos como el número de equipos que pueden alcanzar posiciones destacadas y la incertidumbre en la clasificación general. Aplicamos el modelo a 24 instancias de torneos reales de distintos países, obteniendo mejoras significativas frente a los calendarios históricos. Identificamos desafíos computacionales debidos al rápido crecimiento del problema con el número de equipos, lo que señala

la necesidad de algoritmos de descomposición. Concluimos que considerar dinámicamente el orden de los partidos permite capturar atractivo potencial que los enfoques estáticos tradicionales desaprovechan.

Keywords: programación de torneos , atractivo de torneos , programación entera , fútbol , calendarización dinámica

1. Introducción

La calendarización de un torneo deportivo (asignar partidos a fechas) es un problema clásico de optimización combinatoria. En la práctica se realiza de forma *estática*: el calendario completo se define antes del inicio de la liga y no se modifica durante su desarrollo. Aunque se han incorporado criterios como la equidad en los viajes, la minimización de interrupciones de localía y la maximización de partidos transmitidos por televisión (Ribeiro, 2011), el atractivo del torneo se trata solo de forma indirecta o se ignora.

El atractivo de una competición depende, entre otros factores, de la incertidumbre en los resultados y del número de equipos que mantienen posibilidades reales de alcanzar posiciones relevantes a lo largo de la temporada (Borland & Macdonald, 2003). Un calendario estático puede hacer que varios equipos queden matemáticamente sin incentivos semanas antes del final, reduciendo el interés del torneo.

Trabajos previos han abordado el *diseño* de torneos considerando atraktividad (Devriesere y col., 2025; P. Scarf & Bilbao, 2006; P. A. Scarf & Shi, 2008), pero optimizan el *formato* (cuántas fases, cuántos grupos) y no el orden de los partidos dentro de un formato dado. La calendarización *dinámica* tampoco es nueva: los sorteos progresivos de la Liga de Campeones, el sistema suizo del ajedrez, o el esquema de Dong y col. (2023) para deportes electrónicos usan resultados conocidos para decidir enfrentamientos futuros, pero ninguno aborda una liga *Doble Round-Robin* con enfrentamientos prefijados, donde solo se decide en qué fecha de la segunda rueda juega cada pareja.

Este trabajo responde a la pregunta: ¿cuánto atractivo adicional podría capturar si el orden de los partidos de la segunda rueda se decidiera dinámicamente, conociendo los resultados de la primera? Formulamos el problema SSTPA (*Scheduling Soccer Tournament under measures of Attractiveness*) como un modelo de programación lineal entera y lo comparamos con los calendarios históricos de 24 ligas reales.

2. Descripción del problema y supuestos

Consideramos un torneo *Doble Round-Robin* con conjunto de equipos $\mathcal{I}$, conjunto de partidos $\mathcal{N}$ y conjunto de fechas $\mathcal{F}$. El enfoque es dinámico: la primera

rueda se juega con un calendario prefijado y, al concluir, sus resultados se usan para asignar los partidos de la segunda rueda a las fechas restantes, respetando que cada partido ocurra en exactamente una fecha, cada equipo juegue un partido por fecha, y cada equipo reciba un patrón válido de localías y visitas (con un límite en el número de *breaks*). El modelo opera bajo dos supuestos: *información perfecta* (los resultados de la segunda rueda se conocen de antemano, usando datos históricos reales) e *invarianza del resultado* (el resultado de un partido no depende de la fecha en que se dispute). Estos supuestos permiten calcular el atractivo exacto de cada asignación posible y funcionan como cota superior del atractivo alcanzable por cualquier enfoque dinámico real.

3. Modelo de programación lineal entera

Variables principales. La variable binaria x_{nl} indica si el partido $n \in \mathcal{N}$ se juega en la fecha $l \in \mathcal{F}$. Para evaluar el rango de posiciones de un equipo i en la fecha l no basta con calcular el mejor y peor punteo de i de forma aislada: es necesario construir, para cada rival j , un escenario completo de resultados en *todos* los partidos restantes que favorezca o perjudique a i en esa comparación específica. Para ello se definen las variables binarias v_{ni}^{lf} , a_{ni}^{lf} y e_{ni}^{lf} (y sus análogas $\hat{v}_{ni}^{lf}$, $\hat{a}_{ni}^{lf}$, $\hat{e}_{ni}^{lf}$ para el peor escenario), que indican si el partido n en la fecha $f > l$ se resuelve como victoria local, victoria visitante o empate dentro del escenario optimizado para i . A partir de estas se calcula p_{ji}^{lf} : los puntos que tendría el equipo j al final de la fecha f bajo el escenario de resultados construido para favorecer a i (conociendo la tabla hasta la fecha l). Esta indexación por *par* de equipos (j, i) es la que permite capturar la interacción entre todos los equipos y todos los partidos restantes, en lugar de tratar cada punteo de forma independiente. Las posiciones β_i^l (mejor) y $\hat{\beta}_i^l$ (peor) se obtienen comparando p_{ii}^{lF} (el punteo de i en su propio mejor escenario, fecha final F) contra p_{ji}^{lF} de cada rival, y análogamente con el peor escenario. La variable binaria z_{il}^{uv} toma valor 1 si el equipo i tiene rango (u, v) en la fecha l .

Función objetivo. El *reward* del equipo i en la fecha l con rango (u, v) premia tres aspectos: (i) que el equipo aún pueda alcanzar posiciones con consecuencias deportivas (campeonato, clasificación a copas internacionales o descenso), (ii) la amplitud del rango $v - u$, que refleja la incertidumbre sobre la posición final, y (iii) la proximidad al final del torneo, incentivando que las definiciones más importantes queden para las últimas fechas. Si $u = v$, la posición del equipo ya está determinada y su *reward* es cero. Estos principios se traducen en los siguientes coeficientes: $pp_{uv} = 0$ si $u = v$ y 1 en otro caso (penaliza equipos sin incentivos); $mp_{uv} \geq 1$ es un bono para rangos que incluyen el primer o último puesto; $dp_l = (l - \min \mathcal{F} + 1)^2$ pondera con mayor peso las fechas finales; y ud_u , ld_v miden la distancia de las posiciones u y v al centro de la tabla, premiando posiciones extremas. Formalmente:

$$RF_{uv}^{il} = pp_{uv} \cdot mp_{uv} \cdot dp_l \cdot (w_1 \max(ud_u, ld_v)^2 + w_2 (v - u)^2 + 1), \quad (1)$$

donde $w_1, w_2 \geq 0$ calibran la importancia relativa de cada término. La función objetivo es:

$$\text{máx} \sum_{i \in \mathcal{I}} \sum_{l \in \mathcal{F}} \sum_{u \leq v} RF_{uv}^{il} \cdot z_{il}^{uv}. \quad (2)$$

Restricciones. Cada partido se programa en exactamente una fecha, y cada equipo juega exactamente un partido por fecha:

$$\sum_{l \in \mathcal{F}} x_{nl} = 1 \quad \forall n \in \mathcal{N}, \quad (3)$$

$$\sum_{n: H_n=i \text{ o } A_n=i} x_{nl} = 1 \quad \forall i \in \mathcal{I}, l \in \mathcal{F}. \quad (4)$$

A cada equipo se le asigna un único patrón válido de localías $y_{ip} \in \{0, 1\}$, coherente con el fixture asignado en cada fecha:

$$\sum_{p \in \mathcal{P}} y_{ip} = 1 \quad \forall i \in \mathcal{I}. \quad (5)$$

Para cada partido n jugado en la fecha $f > l$, exactamente uno de los tres resultados posibles ocurre en el escenario construido para i :

$$v_{ni}^{lf} + e_{ni}^{lf} + a_{ni}^{lf} = x_{nl} \quad \forall n \in \mathcal{N}, i \in \mathcal{I}, l, f \in \mathcal{F} : f > l. \quad (6)$$

Los puntos de cada equipo j en el escenario optimizado para i , acumulados hasta la fecha f , suman los puntos ya obtenidos hasta l más los que resultan de los resultados v , a , e asignados a sus partidos entre l y f :

$$p_{ji}^{lf} = pts_j^{0,l} + \sum_{n: H_n=j \text{ o } A_n=j} \sum_{\theta: l < \theta \leq f} 3v_{ni}^{l\theta} + \sum_{n: A_n=j} \sum_{\theta: l < \theta \leq f} 3a_{ni}^{l\theta} + \sum_{n: H_n=j \text{ o } A_n=j} \sum_{\theta: l < \theta \leq f} e_{ni}^{l\theta} \quad (7)$$

para todo $i, j \in \mathcal{I}, l, f \in \mathcal{F} : f > l$, con $pts_j^{0,l}$ los puntos reales de j acumulados hasta la fecha l ; la ecuación análoga con $\hat{v}$, $\hat{a}$, $\hat{e}$ define $\hat{p}_{ji}^{lf}$. Con M una constante grande, $\alpha_{ji}^l \in \{0, 1\}$ indica si j termina con más puntos que i en el mejor escenario de i (fecha final F), y $\hat{\alpha}_{ji}^l$ lo análogo en el peor escenario:

$$p_{ji}^{lF} - p_{ii}^{lF} \leq M(1 - \alpha_{ji}^l) \quad \forall i \neq j, \quad (8)$$

$$\hat{p}_{ii}^{lF} - \hat{p}_{ji}^{lF} \leq M\hat{\alpha}_{ji}^l \quad \forall i \neq j. \quad (9)$$

Finalmente, las posiciones β_i^l y $\hat{\beta}_i^l$ cuentan cuántos equipos superan a i en cada escenario:

$$\beta_i^l = |\mathcal{I}| - \sum_{j \neq i} \alpha_{ji}^l, \quad \hat{\beta}_i^l = 1 + \sum_{j \neq i} (1 - \hat{\alpha}_{ji}^l) \quad \forall i \in \mathcal{I}, l \in \mathcal{F}. \quad (10)$$

Estas posiciones se vinculan con la variable z_{il}^{uv} de la función objetivo: a cada equipo y fecha se le asigna exactamente un rango (u, v) , que debe coincidir con

β_i^l y $\hat{\beta}_i^l$:

$$\sum_{u \leq v} z_{il}^{uv} = 1, \quad \sum_{u \leq v} u \cdot z_{il}^{uv} = \beta_i^l, \quad \sum_{u \leq v} v \cdot z_{il}^{uv} = \hat{\beta}_i^l \quad \forall i \in \mathcal{I}, l \in \mathcal{F}. \quad (11)$$

Las restricciones de cálculo de p_{ji}^{lf} dependen de equipos, partidos y dos índices de fecha, lo que da complejidad $\mathcal{O}(|\mathcal{I}|^5)$ y se traduce en un crecimiento muy rápido del tamaño del problema al aumentar el número de equipos.

4. Resultados

Se trabajó con 24 instancias de ligas reales de once países: Albania, Azerbaiyán, Bosnia y Herzegovina, Chile, Croacia, Grecia, Hungría, Inglaterra, Islandia, Paraguay y Suiza, abarcando desde torneos de 8 equipos hasta torneos de 20 (Premier League inglesa).

Para evaluar el atractivo de cualquier calendario, con independencia de cómo fue construido, se definen cinco métricas: (i) M_1 : promedio, sobre todas las fechas, del número de equipos que aún pueden ser campeones ($\beta_i^l = 1$); (ii) M_2 : ídem para equipos en riesgo de descenso ($\hat{\beta}_i^l = |\mathcal{I}|$); (iii) M_3 : promedio del ancho del rango $\hat{\beta}_i^l - \beta_i^l$; (iv) M_4 : porcentaje de pares (equipo, fecha) con $\beta_i^l = \hat{\beta}_i^l$ (equipo sin incentivos); y (v) M_5 : suma total del *reward* normalizada.

En todas las instancias donde el modelo encontró solución óptima o de buena calidad, el calendario optimizado supera al histórico en las cinco métricas: M_1 y M_2 aumentan (más equipos mantienen opciones durante más fechas), M_4 disminuye (menos pares equipo-fecha sin incentivos), y M_3 y M_5 mejoran consistentemente. Las mejoras son especialmente notorias en torneos pequeños, donde el modelo alcanza la optimalidad. Para torneos de mayor tamaño (Grecia, Chile, Inglaterra), el rápido crecimiento polinomial del problema supera los recursos computacionales disponibles, confirmando la necesidad de métodos de resolución más escalables.

5. Conclusiones y trabajo futuro

El modelo SSTPA demuestra que la calendarización estática desaprovecha atractivo significativo: bajo información perfecta e invarianza de resultados, construye calendarios de la segunda rueda que mejoran todas las métricas de atraktividad frente a los calendarios históricos. Como trabajo futuro, el desafío inmediato es desarrollar esquemas de descomposición (Benders, generación de columnas) que permitan resolver instancias de mayor tamaño, abriendo el camino hacia versiones que no requieran el supuesto de información perfecta.

Referencias

Borland, J. & Macdonald, R. (2003). Demand for sport. *Oxford Review of Economic Policy*, 19(4), 478-502.

- Devriesere, K., Csató, L. & Goossens, D. (2025). Tournament design: A review from an operational research perspective. *European Journal of Operational Research*, 324(1), 1-21.
- Dong, Z.-L., Ribeiro, C. C., Xu, F., Zamora, A., Ma, Y. & Jing, K. (2023). Dynamic scheduling of e-sports tournaments. *Transportation Research Part E: Logistics and Transportation Review*, 169, 102988.
- Ribeiro, C. C. (2011). Sports scheduling: Problems and applications. *International Transactions in Operational Research*, 19(1-2), 201-226.
- Scarf, P. & Bilbao, M. (2006). *The optimal design of sporting contests* (inf. téc. N.º 320/06). Salford Business School, University of Salford.
- Scarf, P. A. & Shi, X. (2008). The importance of a match in a tournament. *Computers & Operations Research*, 35(7), 2406-2418.