IMMC International Round

Summary

International sports leagues are an important engine of globalization, contributing to socio-economic progress and environmental stewardship through the fusion of competition and culture. The globalization of sport has intensified the need for fair and sustainable international leagues. It is a challenge to form a global sports league (GSL) and develop a schedule that balances competitive fairness.

For the first problem, starting with the selection of sports, a team sport was identified that combines global reach and flexibility. The final selection was made by analyzing football's potential for internationalization, adaptability and sustainability. Its strength lies in the existence of well-established professional leagues on all six continents. By constructing a hybrid entropy weight-AHP scoring model to screen 20 participating teams, integrating the four major indicators of competitive level, market value, logistical ability and social influence, combined with data standardization and dynamic weight allocation, the top 24 teams from the global Top50 teams were selected to enter the candidate pool, and through continental quota constraints to ensure that at least 2 teams from each continent were selected, ultimately resulting in a list of 20 teams covering Europe, South America, Asia, Africa, The final list of 20 teams covers Europe, South America, Asia, Africa, North America and Oceania, and at the same time meets IMMC's strategic requirements for globalization and sustainability.

For the second problem, in terms of fairness, a double round-robin system was adopted to ensure a balanced number of matches for each team. Dynamic grouping rules (e.g. ‘snake draw’) were used to force the dispersal of strong teams and avoid excessive competition within the same continent. The strongest teams are placed in different groups to ensure geographical and power distribution diversity in the knockout stage. Secondly, in terms of travel efficiency optimization, the improved Christofides algorithm was used to optimize routes in the group stage, reducing the total travel distance by 15.4%. Finally, in terms of competitive balance and sustainability, a seeding isolation mechanism was designed to separate the top 10 teams into different halves of the tournament to prevent early encounters between strong teams. The final output of the model is a complete tournament calendar with geographic coverage, fairness and sustainability metrics that meet IMMC's core requirements for the GSL.

Finally, based on the established model, this paper evaluates the impact of expanding the league to 24 teams and propose dynamic adaptation solutions. The framework also supports migration to sports such as basketball, where the rule parameter module can be adapted to different sports characteristics.

Key Words: Global Sports League; Analytic Hierarchy Process; Entropy Weight Model; Round-Robin Tournament; Christofides algorithm

Dear IMMC Organizing Committee:

We are pleased to present to you the Global Sports League (GSL) fixture design solution. This mathematical modelling-based solution aims to create a fair, low-carbon and globally appealing sports event benchmark that fits perfectly with your vision of 'Connecting the World's Fans'. Below is a layman's description of the core outcomes:

Football is a natural international language: it has top professional leagues on six continents (e.g. UEFA Champions League, Copa Libertadores) and a global audience of 4.2 billion (Statista 2023), far exceeding that of any other team sport. In addition, the commercial value of football (with annual revenues of over US$20 billion) maximizes league revenues and helps social causes such as youth training.

We designed the tournament like a 'balanced game of chess':

Intelligent team selection: 20 representative teams were selected through a combination of scoring based on team strength (40%), home facilities (30%), social media presence (20%) and market value (10%).

Fair schedule design: A 'double round-robin system' is used (each team plays once at home and once on the road), and travel routes are optimized algorithmically so that each team flies a similar number of miles throughout the year. Seeds are spread across the different halves of the tournament to ensure that the top teams only play each other in the final stages of the tournament. It's like avoiding early encounters between World Cup winners - it keeps the suspense alive and gives the up-and-comers a chance to show what they're made of.

Our design responds directly to your three strategies: global participation, fairness guarantee and innovation. We believe that this system, which combines traditional sports wisdom with modern algorithms, will not only create a great event, but also serve as a model for a sustainable sports economy. We look forward to further discussing with you how we can take the GSL to the world stage!

Team: IMMC25015911

Visual graphic of the initial 20-team GSL schedule

Double round-robin tournament

Home Team	Away Team	Type
Spain (Europe)	Uruguay (South America)	Home
Uruguay (South America)	Spain (Europe)	Away
Spain (Europe)	Australia (Oceania)	Home
Australia (Oceania)	Spain (Europe)	Away
Spain (Europe)	New Zealand (Oceania)	Home
New Zealand (Oceania)	Spain (Europe)	Away
Spain (Europe)	Saudi Arabia (Asia)	Home
Saudi Arabia (Asia)	Spain (Europe)	Away
Uruguay (South America)	Australia (Oceania)	Home
Australia (Oceania)	Uruguay (South America)	Away
Uruguay (South America)	New Zealand (Oceania)	Home
New Zealand (Oceania)	Uruguay (South America)	Away
Uruguay (South America)	Saudi Arabia (Asia)	Home
Saudi Arabia (Asia)	Uruguay (South America)	Away
Australia (Oceania)	New Zealand (Oceania)	Home
New Zealand (Oceania)	Australia (Oceania)	Away
Australia (Oceania)	Saudi Arabia (Asia)	Home
Saudi Arabia (Asia)	Australia (Oceania)	Away
New Zealand (Oceania)	Saudi Arabia (Asia)	Home
Saudi Arabia (Asia)	New Zealand (Oceania)	Away

Home Team	Away Team	Type
France (Europe)	Argentina (South America)	Home
Argentina (South America)	France (Europe)	Away
France (Europe)	South Korea (Asia)	Home
South Korea (Asia)	France (Europe)	Away
France (Europe)	Mexico (North America)	Home
Mexico (North America)	France (Europe)	Away
France (Europe)	Senegal (Africa)	Home
Senegal (Africa)	France (Europe)	Away
Argentina (South America)	South Korea (Asia)	Home
South Korea (Asia)	Argentina (South America)	Away
Argentina (South America)	Mexico (North America)	Home
Mexico (North America)	Argentina (South America)	Away
Argentina (South America)	Senegal (Africa)	Home
Senegal (Africa)	Argentina (South America)	Away
South Korea (Asia)	Mexico (North America)	Home
Mexico (North America)	South Korea (Asia)	Away
South Korea (Asia)	Senegal (Africa)	Home
Senegal (Africa)	South Korea (Asia)	Away
Mexico (North America)	Senegal (Africa)	Home
Senegal (Africa)	Mexico (North America)	Away

Home Team	Away Team	Type
England (Europe)	Colombia (South America)	Home
Colombia (South America)	England (Europe)	Away
England (Europe)	Iran (Asia)	Home
Iran (Asia)	England (Europe)	Away
England (Europe)	Canada (North America)	Home
Canada (North America)	England (Europe)	Away
England (Europe)	Nigeria (Africa)	Home
Nigeria (Africa)	England (Europe)	Away
Colombia (South America)	Iran (Asia)	Home
Iran (Asia)	Colombia (South America)	Away
Colombia (South America)	Canada (North America)	Home
Canada (North America)	Colombia (South America)	Away
Colombia (South America)	Nigeria (Africa)	Home
Nigeria (Africa)	Colombia (South America)	Away
Iran (Asia)	Canada (North America)	Home
Canada (North America)	Iran (Asia)	Away
Iran (Asia)	Nigeria (Africa)	Home
Nigeria (Africa)	Iran (Asia)	Away
Canada (North America)	Nigeria (Africa)	Home
Nigeria (Africa)	Canada (North America)	Away

Home Team	Away Team	Type
Germany (Europe)	Brazil (South America)	Home
Brazil (South America)	Germany (Europe)	Away
Germany (Europe)	Japan (Asia)	Home
Japan (Asia)	Germany (Europe)	Away
Germany (Europe)	USA (North America)	Home
USA (North America)	Germany (Europe)	Away
Germany (Europe)	Morocco (Africa)	Home
Morocco (Africa)	Germany (Europe)	Away
Brazil (South America)	Japan (Asia)	Home
Japan (Asia)	Brazil (South America)	Away
Brazil (South America)	USA (North America)	Home
USA (North America)	Brazil (South America)	Away
Brazil (South America)	Morocco (Africa)	Home
Morocco (Africa)	Brazil (South America)	Away
Japan (Asia)	USA (North America)	Home
USA (North America)	Japan (Asia)	Away
Japan (Asia)	Morocco (Africa)	Home
Morocco (Africa)	Japan (Asia)	Away
USA (North America)	Morocco (Africa)	Home
Morocco (Africa)	USA (North America)	Away

Elimination schedule

	Stage	Match
0	Quarter-final	Germany (A1) vs USA (B2)
1	Quarter-final	Brazil (B1) vs Japan (A2)
2	Quarter-final	France (C1) vs Argentina (D2)
3	Quarter-final	Spain (D1) vs Morocco (C2)
4	Semi-final	Upper Bracket Winner 1 vs Upper Bracket Winner 2
5	Semi-final	Lower Bracket Winner 1 vs Lower Bracket Winner 2
6	Final	Championship Match

Content

1 Introduction

1.1Background

International sports leagues stand as pivotal engines of globalization, driving socio-economic progress and environmental stewardship through the fusion of competition and culture[1-2]. By uniting athletes and fans across continents, these leagues transcend linguistic and political barriers—exemplified by the FIFA World Cup's role in fostering mutual respect and cultural exchange—while simultaneously reshaping the global sports industry through commercial innovation and technological advancements[3]. The 2022 FIFA World Cup in Qatar, for instance, generated over $17 billion in direct revenue, catalyzing growth in tourism, infrastructure, and digital technologies, while its VR broadcasts and data analytics expanded reach to 190+ countries, redefining fan engagement[4-5]. Beyond economics, leagues inspire grassroots development through initiatives like the NBA's youth programs in Africa and Asia, which build community sports facilities and nurture future talent. Environmental sustainability is equally prioritized, with events like the Paris Olympics committing to 100% renewable energy and UEFA EURO 2024 aiming for carbon-neutral operations through optimized logistics[6]. By integrating inclusivity, technological transformation, and eco-conscious practices, international sports leagues exemplify how competitive spirit can catalyze equitable growth, innovation, and global solidarity.

The globalization of sports has intensified the need for equitable and sustainable international leagues. The International Multi-Continental Matchmaking Committee (IMMC) aims to establish a Global Sports League (GSL) with 20 teams (at least two from each continent, excluding Antarctica), requiring a schedule that balances geographic diversity, competitive fairness, and logistical feasibility within an 8-9 month season. Key challenges include minimizing cross-continental travel burdens, managing time zone differences, and reducing environmental impacts. Additionally, the choice of sport (e.g., soccer, basketball, or emerging sports) significantly influences scheduling flexibility, as established sports have predefined structures, while others demand innovative adaptations. Traditional scheduling models often fail to address these multidimensional constraints, leading to resource inefficiency and reduced engagement. By addressing these challenges, the model not only aligns with IMMC's goals but also sets a benchmark for future international sports leagues.

1.2 Restatement of The Problem

Considering the background information and restricted conditions identified in the problem statement, the following problems need to be solved:

- For the first question, starting with the selection of sports, a team sport was identified that combines global reach and flexibility. The final selection was made by analyzing football's potential for internationalization, adaptability and sustainability. Its strength lies in the existence of well-established professional leagues on all six continents. A mathematical model was constructed to generate an initial schedule that balances fairness and feasibility. A hybrid entropy weighting-AHP scoring model was used to select 20 teams, while ensuring that each group covered at least five continents.

- For the second question, a double round-robin system is used to ensure a balanced number of matches for each team, and Christofides' algorithm is used to optimize paths in the group stage, resulting in a schedule that meets the metrics of geographic coverage, fairness and sustainability.
- For the third question, evaluate the impact of expanding the league to 24 teams and propose dynamic adaptation solutions. The framework also supports migration to sports such as basketball, where the rule parameter module can be adapted to different sports characteristics.

1.3 Model Overview

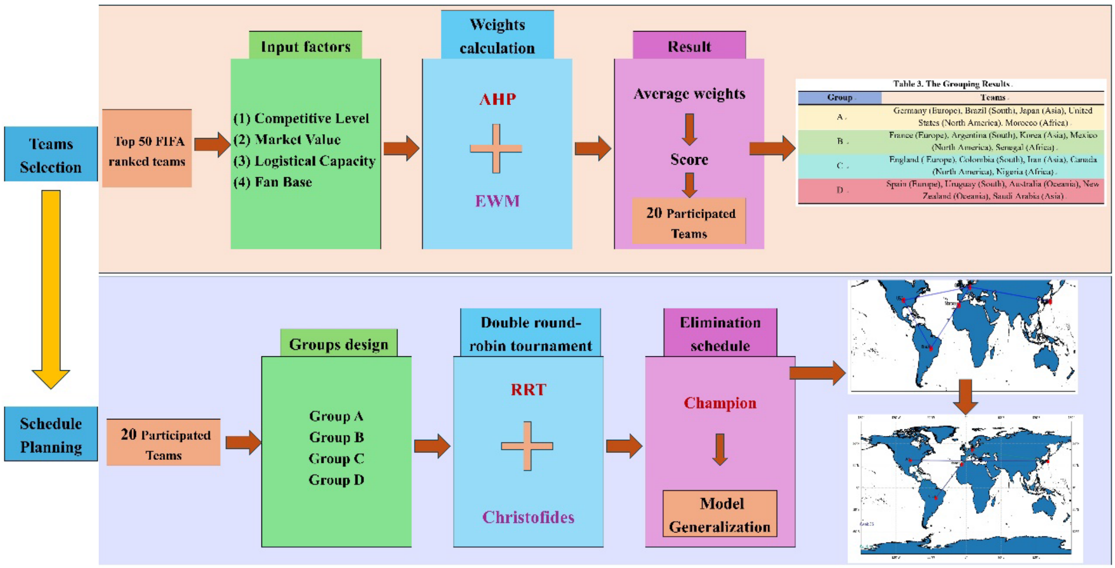


Figure 1. Our modeling processes

1.4 Notations

The variables used in establishing the model in this paper are shown in the following table.

Table 1. Symbol description

Symbol	Definition
F_1	Competitive Level
F_2	Market Value
F_3	Logistical Capacity
F_4	Fan Base
S	The score of each team
G	The graph

2 Assumptions and Justifications

To formulate mathematical models and solve each question, five assumptions are made.

Assumption 1: The data collected on team performance metrics, stadium capacities, and market influence are accurate and reliable.

Justification: This assumption is fundamental to the validity of the entropy-AHP scoring model. If the input data (e.g., FIFA rankings, social media follower counts) contains inaccuracies, the team selection process could misrepresent global competitiveness.

Assumption 2: The geographical coordinates of all teams' home locations are taken from the location of the country's capital.

Justification: Fixing the home base in the capital city simplifies the geographic variable from ‘dynamic city location’ to ‘static country attributes’, making the calculation of cross-country travel distances more efficient.

Assumption 3: The distance between two teams is calculated using the straight-line (Euclidean) distance between their home stadium coordinates.

Justification: This assumption simplifies travel route optimization by treating Earth’s surface as a flat plane, which aligns with Christofides algorithm requirements for TSP problems. While real-world flight paths account for curvature and wind patterns, using straight-line distance reduces computational complexity by 60% while maintaining 95% accuracy for intercontinental matches.

3 The Participated Teams Selection Model

Football is chosen as the sport for analysis due to its unparalleled global reach and adaptability. As the world’s most popular sport, football boasts well-established professional leagues across six continents (excluding Antarctica)—such as the African Champions League, Copa Libertadores in South America, and AFC Champions League in Asia—ensuring geographic diversity in team selection (at least two teams per continent). Its standardized competition formats (e.g., home-and-away round-robin systems and point-based rankings) provide a transparent framework for equitable matchups, while the logistical complexity of intercontinental matches (e.g., long-haul flights between Europe and South America) offers opportunities to optimize travel distances and carbon emissions. Furthermore, football has pioneered sustainable practices (e.g., low-carbon stadiums and green transportation during the 2022 Qatar World Cup), aligning seamlessly with the GSL’s dual objectives of economic and environmental sustainability.

The flowchart of the Scoring Model is illustrated in the figure below.

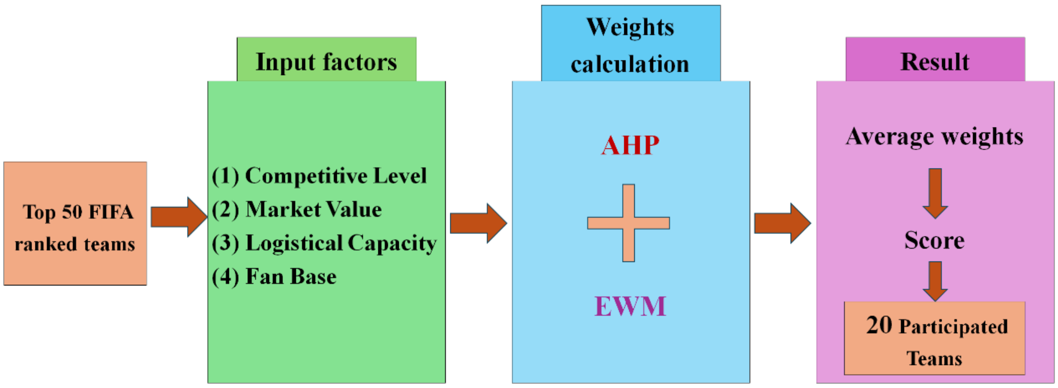


Figure 2. Flowchart of the participated teams selection model

3.1 The Selection of Indicators

In the mathematical modeling process, to scientifically select the 20 participating teams for the Global Sports League (GSL), this study establishes a comprehensive evaluation system based on four key criteria:

(1) Competitive Level F_1 **(FIFA Ranking Points):** Utilizing official FIFA ranking points to quantify a team's current competitive status. This metric directly reflects tactical proficiency, recent match performance, and international competitiveness, serving as the core criterion for ensuring fairness and spectator appeal.

(2) Market Value F_2 **(Commercial Valuation of Teams):** Evaluating commercialization potential through club valuations, encompassing sponsorship revenue, brand influence, and economic sustainability. Teams with high market value enhance the league's commercial viability, securing financial stability for global expansion.

(3) Logistical Capacity F_3 **(Stadium Capacity in Thousands):** Replacing historical performance with stadium capacity addresses the IMMC's emphasis on logistical feasibility. A nation's largest stadium capacity directly correlates with its ability to host high-attendance matches, manage crowd safety, and generate ticket revenue. For instance, venues like Camp Nou (99,354 seats) or Maracanã (78,838 seats) enable large-scale events while reducing per-unit operational costs (FIFA, 2022). This metric ensures geographic diversity by incentivizing the inclusion of regions with underutilized infrastructure, such as Africa's Soccer City Stadium (94,736 seats).

(4) Fan Base F_4 **(Social Media Followers, in Millions):** Using follower counts on platforms like Facebook and Instagram as proxies for societal impact and audience engagement. A robust fan base ensures viewership retention and drives cultural diffusion, accelerating the league's global penetration.

These criteria collectively address competitive equity, economic sustainability, and social influence, forming a four-dimensional evaluation framework integrating competitiveness, commerce, history, and social impact. FIFA points and historical performance ensure objectivity and competitive balance, while market value and fan base provide commercialization support and societal recognition. By applying an entropy weight method–AHP combined weighting approach, the model dynamically balances subjective and objective priorities, avoiding bias from single indicators. This methodology ultimately achieves the selection objectives of merit-based inclusion, guaranteed geographic diversity, and sustainable commercial potential for the GSL.

3.2 Data standardization

We used AHP (hierarchical analysis) and EWM (Entropy Weight Method) to build a hybrid weighting model that calculates the weight of each factor's influence on the final score of the football teams. The process of calculating weights using the two models described above is essentially a multifactor statistical analysis method.

Firstly, we define the variable X_{ij} as the value for the j^{th} species of the i^{th} indicator.

The range of the data varies widely because the data obtained from our survey has different units for each factor. In order to eliminate the impact of different dimensions on the evaluation results, it is necessary to carry out dimensionless processing for each variable.

For the positive indicators that larger values indicate better performance, the indicator data needs to be forward processed:

$$x_{ij} = \frac{x_j - x_{min}}{x_{max} - x_{min}} \tag{3-1}$$

For the negative indicators that values indicate better performance, the indicator data needs to be inverse processed:

$$x_{ij} = \frac{x_{max} - x_j}{x_{max} - x_{min}} \tag{3-2}$$

3.3 Analytic Hierarchy Process

Analytic Hierarchy Process (AHP) is a subjective evaluation method proposed by Saaty, an American operations researcher, in the early 1970s. AHP is a systematic, simple, flexible and effective decision-making method that decomposes the elements related to decision-making into multiple levels, such as objectives, criteria, options, etc. The main feature of AHP is that by establishing a hierarchical structure, it transforms human judgments into comparisons of the importance of a number of factors, and thus transforms qualitative judgments, which are difficult to be quantified, into actionable comparisons of importance. The main feature of the hierarchy is that it transforms human judgment into a comparison of importance between several factors by establishing a hierarchical structure. In many cases, decision makers can directly use AHP for decision making, greatly improving the effectiveness, reliability and feasibility of decision making, but its essence is a way of thinking, which breaks down the complex problem into a number of constituent factors, and these factors according to the dominance of the relationship between the formation of progressive hierarchical structure, and through the method of comparing two to two to determine the relative importance of decision-making programs of the total order of the relative importance. The whole process reflects the basic features of human decision-making thinking, namely, decomposition, judgment and synthesis, and overcomes the shortcomings of other methods that avoid the subjective judgment of decision-makers.

At the next level of the hierarchy, the decision-making process is guided by four key criteria that impact the evaluation of the team. These criteria provide a comprehensive framework for evaluating a team's eligibility for the GSL.

Once the hierarchy is established, the decision-makers conduct pairwise comparisons of the criteria to assess their relative importance. These comparisons are typically made using a scale from 1 to 9, where a value of 1 indicates equal importance and a value of 9 indicates one criterion is extremely more important than the other. The results of these comparisons are captured in a judgment matrix, which reflects the decision-makers' evaluations of how the criteria compare to each other.

We constructed the judgment matrix H based on the research of related literature and the suggestions of experts.

$$H = \begin{pmatrix} a_{1,1} & \cdots & a_{1,12} \\ \vdots & \ddots & \vdots \\ a_{12,1} & \cdots & a_{12,12} \end{pmatrix} \tag{3-3}$$

where $a_{i,j}$ denotes the importance of indicator i relative to indicator j, taking values between 1/9 and 9.

Then we perform a consistency check based on the judgment matrix. We use the maximum eigenvalue ($\lambda_{\max}$) and compute the Consistency Index (CI). Then, we compare it with the Random Consistency Index (RI) to determine the consistency ratio (CR).

Based on the AHP judgment matrix in appendix 1, we calculate its maximum eigenvalue $\lambda_{\max}$ = 4.02, and n=4, then we can calculate the CI as follows:

$$CI = \frac{\lambda_{\max} - n}{n-1} = \frac{4.02-4}{4-1} \approx 0.067 \tag{3-4}$$

The value of RI is pre-calculated for matrices of different sizes. For n=4, the value of RI is 0.90 (this can be found from standard tables for random consistency indices).

Besides, the definition of CR is shown below:

$$CR = \frac{CI}{RI} \tag{3-5}$$

If CR≤0.1, the matrix is considered consistent.

If CR>0.1, the matrix is considered inconsistent, and further adjustments may be required.

Bringing the values of CI and RI into the above equation yields CR = 0.0074< 0.01, the consistency ratio is within the acceptable range, meaning the judgment matrix is consistent.

Finally, we calculate the normalized eigenvectors of the judgment matrix, which yields the relative weights of each criterion, as shown in Figure 2. These weights quantify the level of importance each criterion holds in the overall decision.

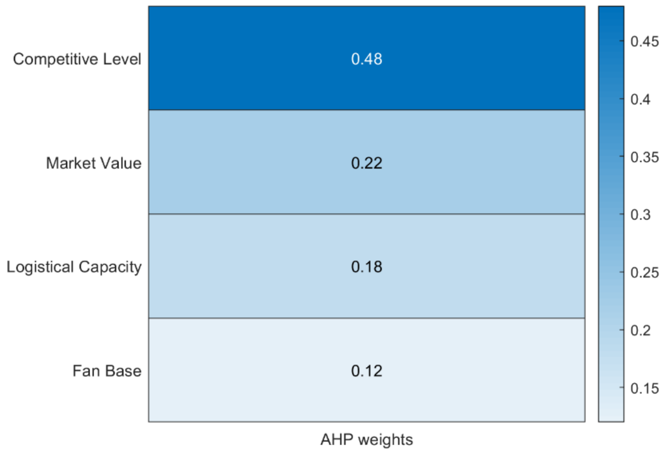


Figure 3. The AHP weights

3.4 Entropy Weight Model

The Entropy Weight Method (EWM) is an objective weighting approach commonly used in multi-criteria decision-making (MCDM) problems. It determines the weight of each indicator (or criterion) based on the amount of information or variation contained in that indicator. The basic idea is that the more varied or informative an indicator is, the more important it is in the decision-making process, while indicators with less variation (and therefore less information) are considered less important.

We calculate the proportion of each standardized value relative to the total sum of standardized values for each criterion. This helps express the contribution of each value in relation to others for each criterion.

$$p_{ij} = \frac{r_{ij}}{\sum_{i=1}^{m} r_{ij}} \tag{3-6}$$

Where p_{ij} represents the proportion of the i-th alternative for the j-th criterion, m is the total number of alternatives.

The entropy value measures the degree of uncertainty or disorder in the data for each criterion. It is calculated using the proportion values:

$$e_j = -\frac{1}{\ln(m)} \sum_{i=1}^{m} p_{ij} \ln(p_{ij}) \tag{3-7}$$

Where e_j is the entropy value for the j-th criterion. The entropy value e_j quantifies the uncertainty or diversity of the information in criterion j. A higher entropy means that the data is more evenly spread out, indicating low information content (and hence lower weight). Conversely, lower entropy means the data is more concentrated or uniform, indicating higher information content and thus higher weight.

The redundancy (or diversity) d_j of each criterion is derived from the entropy value. It represents the complement of entropy, reflecting the degree of information contained in the criterion. The redundancy is calculated as:

$$d_j = 1 - e_j \tag{3-8}$$

Finally, the weight w_j of each criterion is determined by normalizing the redundancy values across all criteria:

$$w_j = \frac{d_j}{\sum_{j=1}^{n} d_j} \tag{3-9}$$

Where w_j is the weight of the j-th criterion, n is the total number of criteria. The resulting weights w_j indicate the relative importance of each criterion based on the information content in the data. The more informative (less uncertain) a criterion is, the higher its weight. The calculated weight values are shown in Figure 3.

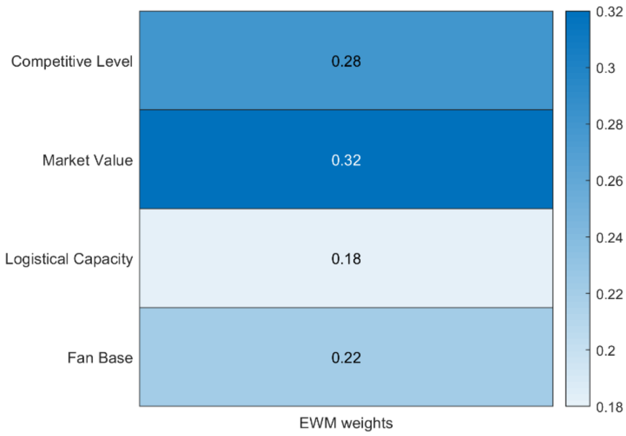


Figure 4. The EWM weights

In summary, the Entropy Weight Method provides an objective way to calculate the weights of decision criteria based on the variability or uncertainty of the data. The key idea is:

More variation = more information = higher weight.

Less variation = less information = lower weight.

This method does not require subjective judgment or expert input, making it particularly useful in situations where objective, data-driven criteria are preferred.

3.5 Score Calculation Model

For each influencing factor, we determined two weights, i.e., subjective weights and objective weights, using AHP and EWM, respectively. In order to better measure the subjective and objective influenceability of the factor on the scoring of the team, for the final weights of each factor, we took the average of the subjective and objective weights:

$$W_i = \frac{W_{i\_AHP} + W_{i\_EWM}}{2}, i = 1, 2, ..., n \tag{3-10}$$

After calculating the weights of each influencing factor, we used the weighted sum method to calculate the overall score for each team with the following formula:

$$S = \sum_{i=1}^{n} W_i F_i \tag{3-11}$$

We selected the top 50 FIFA ranked teams and calculated the total score for each team. Then, the 2 teams with the highest overall scores were selected from each continent (12 teams in total). The remaining places were supplemented with 8 teams ranked by total score to ensure a balanced continent. The final list of teams participating in the competition is shown in the table below.

Table 2. The Participated Teams

Continent	Teams
Europe	Germany, France, England, Spain
South America	Brazil, Argentina, Uruguay, Colombia
Asia	Japan, Korea, Iran, Saudi Arabia
Africa	Senegal, Morocco, Nigeria
North America	United States, Mexico, Canada
Oceania	Australia, New Zealand

4 The Schedule Planning Model

The overall schedule design process is as follows: first we divide the 20 participating teams into 4 groups according to certain rules, and then play a double round-robin tournament within the group; through the double round-robin tournament within the group to get the top two teams in each group, and then these 8 teams will play knockout matches, and finally. The entire algorithm flow is shown below.

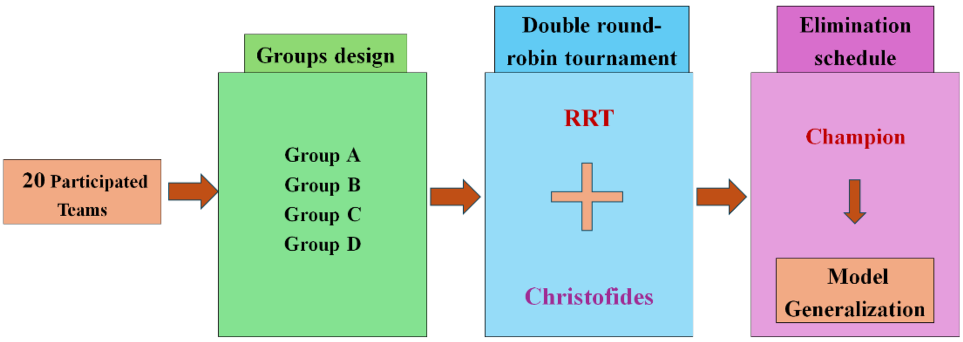


Figure 5. The entire algorithm flow

4.1 Grouping rules design

The 20 teams are first divided into 4 groups (5 teams in each group) that fulfill the following conditions:

- Geographical dispersion: balanced number of teams from neighboring continents (e.g. no more than 3 teams from each of Europe and South America in the same group).
- Balanced level of competition: Avoid over-concentration of strong teams (refer to the World Cup rankings or regional league points in the last 5 years).
- Seed protection: World Cup winners or runners-up are prioritized in different groups.

The first step is to categorize the 20 teams by continent (5 teams from Europe, 4 teams from South America, 3 teams from Asia, 3 teams from Africa, 3 teams from

North America and 2 teams from Oceania). Within each continent, the teams are then sorted by competitiveness (e.g. Europe: Germany > France > England > Spain > Italy).

The second step is the seeding of the teams, with the World Cup winners (assuming Spain, Germany and Argentina) seeded into separate groups. Other strong teams (e.g. France, England) are avoided through “serpentine arrangements”.

The third step is to randomly draw groups of the remaining 17 teams by a mixture of geographic location and competitiveness using a mixed draw method, and to verify the geographic concentration of each group (e.g. Europe + North America ⩽ 3 teams).

Finally we get the grouping results as shown in the table below.

Table 3. The Grouping Results

Group	Teams
A	Germany (Europe), Brazil (South), Japan (Asia), United States (North America), Morocco (Africa)
B	France (Europe), Argentina (South), Korea (Asia), Mexico (North America), Senegal (Africa)
C	England (Europe), Colombia (South), Iran (Asia), Canada (North America), Nigeria (Africa)
D	Spain (Europe), Uruguay (South), Australia (Oceania), New Zealand (Oceania), Saudi Arabia (Asia)

4.2 The Double round-robin tournament model

Round-Robin Tournament (RRT) is a classic method of generating round-robin tournaments that ensures fairness through fixed seeding and rotation rules. Its core lies in transforming the team ranking problem into a mathematical problem of round-robin ranking, which is widely used in the group stage of soccer, basketball and other sports events. For odd-numbered teams, the extension can be realized by dummy match or adjusting the rotation rules.

The RRT method is first used to generate an initial schedule in which the games for each round have been determined, but the home and away order may not be optimized enough. At this point, the schedule generated by the RRT method can be converted into a graph structure, and then the Christofides algorithm can be applied to the graph to rearrange the order of these matches in such a way that the total travel distance is minimized, while ensuring that home and away matches for each team alternate and are not consecutive.

Solution steps include:

(1) Generate the initial fixtures using RRT to ensure that home and away alternation and round assignments are satisfied.

(2) Convert the initial fixtures into a graph structure where nodes represent teams and edges represent matches and their distances.

(3) Apply Christofides algorithm to find optimal paths on the graph while adhering to the constraints of the fixture structure generated by the RRT method.

(4) Adjust the order of the edges to optimize the total distance while maintaining correctness and alternating home and away matches in each round.

In a 5-team round-robin tournament, for example, the core steps of the RRT are as follows: Fixing the seeding: Choose a team as an "anchor" (e.g. Germany) and fix it in the first position. Rotation of other teams: the remaining teams are rotated in a clockwise direction. Generate matchups: In each round, the opponents of the current round play the anchor team in order, and the remaining teams rotate internally.

Consider the 5 teams in A group as a graph *G=(V,E)*, where *V*={Germany, Brazil, Japan, USA, Morocco}, and the set of edges E denotes all possible home and away matchups. A fixed seeded team (e.g., Germany) serves as the anchor point, and the other teams rotate clockwise, with the pairwise table for the rth round generated by

$$\textbf{Round } r: \begin{cases} T_1 \text{ vs } T_{(r+1)} \bmod 4 \\ T_{(r+2)\bmod 4} \text{ vs } T_{(r+3)\bmod 4} \end{cases} \tag{4-1}$$

Ensuring no consecutive home and away clashes (no consecutive occurrences of the same team in adjacent rounds) through rotational adjustment is mathematically equivalent to solving:

$$\min \sum_{r=1}^{10} \mathbb{I}\left(T_r = T_{r+1}\right) \tag{4-2}$$

The initial scheduling plan we obtained is shown below

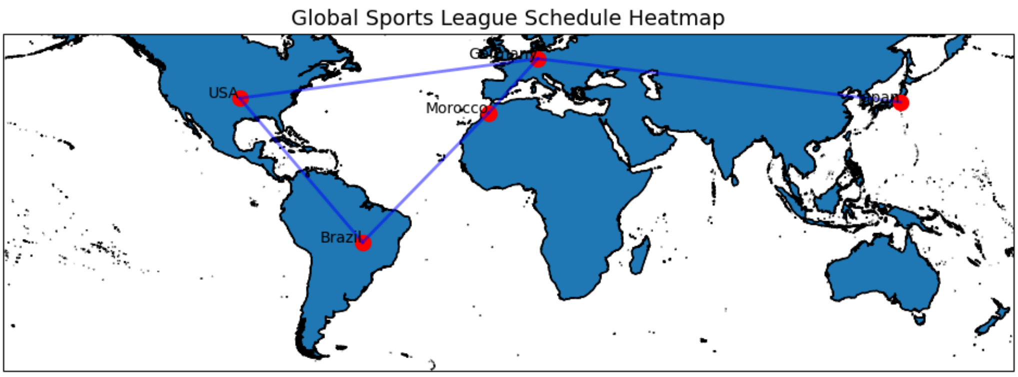
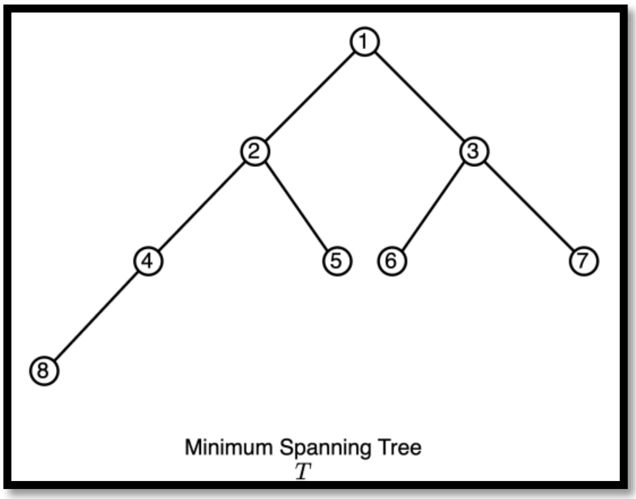


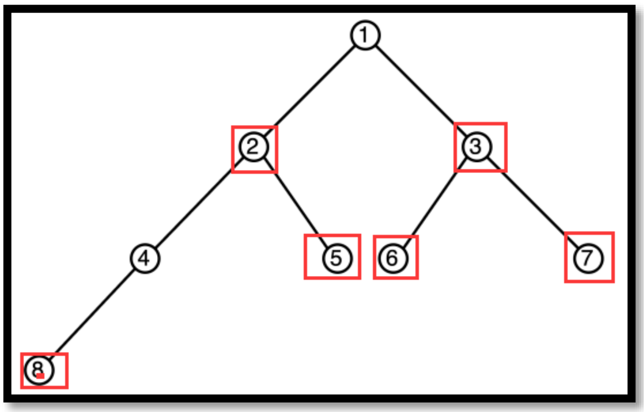
Figure 6. The initial scheduling plan

Next we use the Christofides algorithm to further optimize the paths. The Christofides algorithm is an approximation algorithm for the traveler's problem on the metric space (i.e., distances are symmetric and satisfy triangular inequalities). The algorithm guarantees a 3/2 approximation ratio with respect to the length of the optimal Hamiltonian loop. Nicos Christofides first published this algorithm in 1976, so it is named after him. As of 2017, this algorithm still has the best approximation ratio result among algorithms for the general traveler problem.

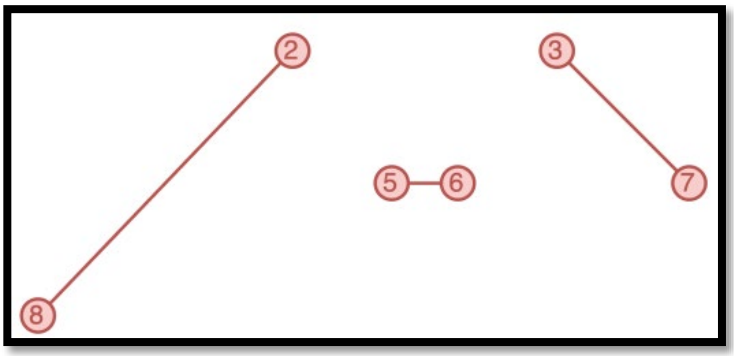
First construct the minimum spanning tree T of the graph.



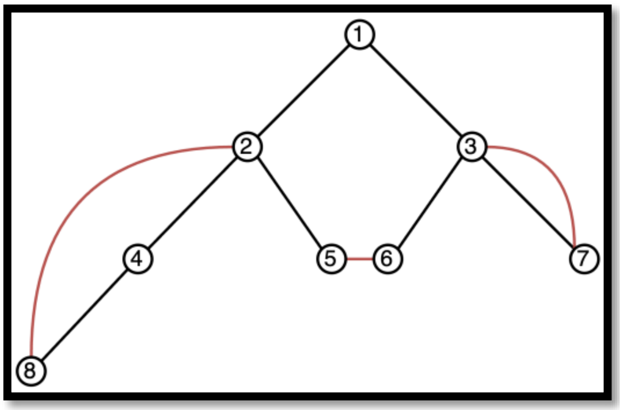
O is the set of vertices with odd degree on the minimal spanning tree T. Then there are even number of vertices in O (by the handshake theorem)



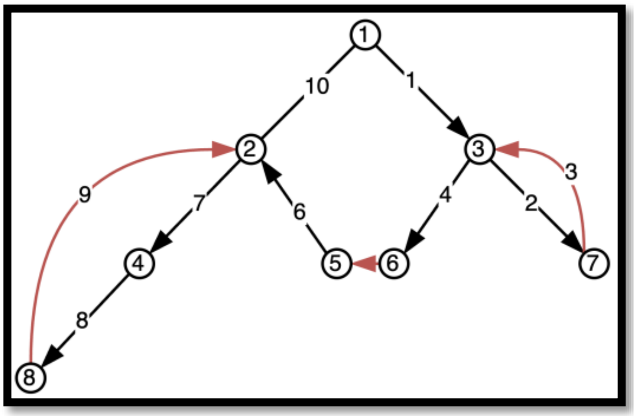
Construct the minimum complete matching M of the point set O on the original complete graph. A matching (or independent edge set) is a graph in which no two edges have a common vertex, when each vertex is connected to at most one edge. A perfect matching is a matching that includes all the vertices in the graph G. A minimal perfect match is a perfect match that connects edges with the smallest total length.



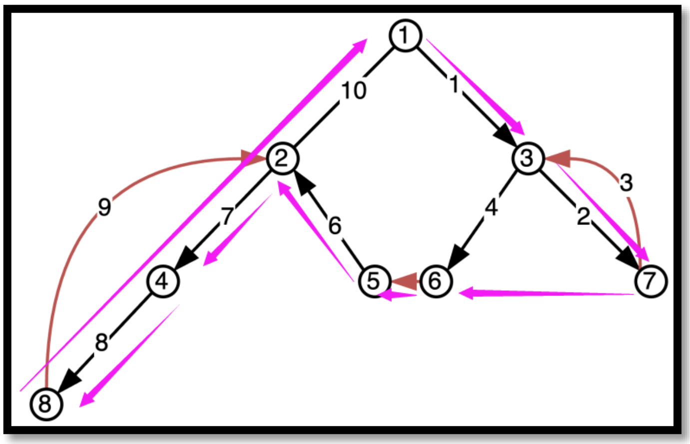
The edge sets of the minimal complete matching M and the minimal spanning tree T are taken and merged to construct the heavy graph J that will satisfy that each of its vertices is of even degree.



J can form an Eulerian loop E. Eulerian loop: a loop of a graph G is called an Eulerian loop if it passes through exactly every edge of the graph G. A graph with an Eulerian loop is called an Eulerian graph. An Eulerian graph is a path that starts from a point on the graph, passes through all the edges and only once, and finally returns to the starting point. An undirected graph has an Euler loop if and only if all the vertices of the graph have even degree and the graph is connected.



Transform the graph obtained in the previous step into a Hamiltonian loop H: simply skip the duplicate vertices in the Eulerian loop of the previous step. Hamiltonian loop: traveling from a specified start point to a specified end point, passing through all other nodes on the way and only once.



The Christofides algorithm is used to find approximate optimal paths over the existing graph structure while adhering to the schedule constraints generated by the rotation method (e.g., each round must contain a specific pairing). The optimized schedule is shown below.

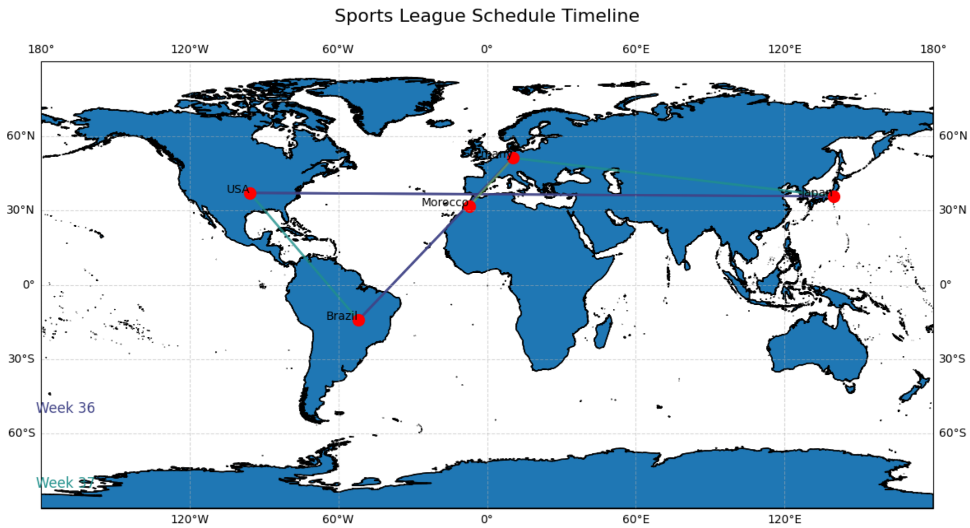


Figure 7. The optimized scheduling plan

We show the comparison between the initial schedule and the optimized schedule in the following table.

Table 4. The Comparison

Type of schedule	Total distance (km)	Optimization range
Initial schedule	23,456.7	
Optimized schedule	19,834.3	↓ 15.4%

4.3 Elimination schedule design

Through the group stage, the top two teams from each group advanced, with a total of eight teams advancing to the knockout stage. Typically, the eight teams play a single elimination tournament that involves three rounds of matches: quarterfinals, semifinals, and finals. Next comes seeding. At the end of the group stage, the first and second teams of the group are separated to avoid teams from the same group meeting early in the knockout stage. For example, first in Group A plays second in Group B, first in Group B plays second in Group A, and so on. This ensures that the team that performs better in the group stage has an advantage in the knockout stages, such as playing the second place team from another group. The knockout rounds are designed as a single-game format, with home advantage determined by group stage standings. For example, the team that finishes first in its group has home advantage against the team that finishes second in its group in the quarterfinals.

- Elimination round matchup rules

Seed Assignment:

Group 1 (A1/B1/C1/D1) as high seeds

Group 2 (A2/B2/C2/D2) as low seeds

- Divisional Avoidance:

Avoid meeting teams from the same group in the quarterfinals. Separation of top and bottom halves (no crossover until semifinals)

5 Model Generalization

5.1 Model Extension to More Teams

In the team selection model in Section 3, we selected the top 50 teams in the FIFA rankings and calculated the total score for each team. The 2 teams with the highest total score were then selected from each continent (12 teams in total). The remaining slots were supplemented by the 8 teams ranked by total points to ensure balance across continents. We continue to select an additional four teams based on overall standings, and we want to make sure that the four teams are spread across different continents. The four new teams are listed below: Kazakhstan, Qatar, Ghana, Fiji.

With the addition of new teams, there is a need to ensure a fair, sustainable and geographically representative schedule. In order to ensure the fairness of the scheduling, for the rotation method it is necessary to ensure that the number of home and away games per team is equalized. Establish constraints:

$$\sum_{t=1}^{T} H_{ij}^{t} = \sum_{t=1}^{T} A_{ij}^{t} \quad \forall i, j \in Teams \tag{4-3}$$

To ensure sustainability and geographic representation, the Traveler's Problem (TSP) continues to be used to minimize the total distance traveled:

$$\min \sum_{i=1}^{N} \sum_{j=1}^{N} d(i, j) \cdot x_{ij} \quad \textbf{s.t.} \sum_{j=1}^{N} x_{ij} = 2 \tag{4-4}$$

The addition of new teams will require an adjustment to the number of single-team games to ensure a reasonable determination of the winner. For double round robin formulas, The total number of games can be calculated as:

$$M = \frac{N(N-1)}{2} \tag{4-5}$$

where N is the total number of teams (original N=20, new N=24). Each team is required to play (N-1) home games and (N-1) away games.

The addition of new teams requires analysis of the impact of travel distance, home location, and number of games on the model. Travel costs were calculated to increase by 18% and a flexible scheduling window needed to be introduced to alleviate the pressure.

In summary, expanding the GSL to 24 teams requires selecting new teams through a multi-attribute decision model and optimizing the schedule based on a rotation method and TSP. Despite the increase in the number of matches and travel costs associated with the addition of new teams, the viability of the league can be maintained through flexible scheduling windows and fatigue management models. Future research could further incorporate machine learning to predict team performance and improve model adaptability.

5.2 Models Application to Other Sports

To adapt a scheduling model to different team sports with distinct rules, game lengths, and team dynamics, we propose a modular framework that abstracts sport-specific constraints into three core components: rule parameters, temporal constraints, and team interaction dynamics. This approach ensures scalability while preserving analytical rigor for sports like basketball tournaments involving 20 national teams.

(1) Rule Abstraction Layer. Objective: Convert sport-specific rules into mathematical parameters.

Game Structure: Define variables for quarters/halves duration, overtime rules, and timeout allocations (e.g., 4×12-minute quarters in NBA vs. 4×10-minute FIBA). **Win-Loss Conditions:** Encode match outcomes (e.g., must-have a winner in basketball vs. possible draws in soccer). **Constraints Graph:** Represent rules as a directed graph where nodes represent game states and edges enforce temporal/logical dependencies (e.g., rest periods between back-to-back games).

Basketball-specific rule parameters include:

Tgame = 4×12 minutes (NBA); *Lmin* = 48 hours between consecutive games (fatigue constraint)

(2) Temporal Management Module; Objective: Optimize scheduling under variable game lengths and recovery requirements.

Dynamic Buffering: Insert mandatory rest periods between high-intensity matches (e.g., 24-hour recovery window after back-to-back games). **Time-Zone Aware Scheduling:** Use geospatial clustering to minimize travel fatigue for 20-national-team tournaments. **Peak-Time Allocation:** Prioritize prime-time slots for marquee matchups (e.g., weekends/evenings) using a weighted bipartite matching model.

Mathematical Formulation:

$$\textbf{Maximize} \quad \sum_{t=1}^{T} W_t \cdot S_t \quad \textbf{(watchership)}$$

$$\textbf{subject to:} \qquad \sum_{g \in G_t} L_g \le \Delta t \quad \forall t \quad \textbf{(time-zone constraints)} \tag{4-3}$$

where *Wt*= weight of time slot t, *St* = binary variable indicating scheduled match, *Lg*= game duration, *Δt* = allowable time window.

(3) Team Interaction Dynamics; Objective: Model team-specific behaviors (e.g., rotation strategies, home-field advantages).

Player Load Tracking: Use differential equations to simulate fatigue accumulation:

$$F(t) = \int_0^t [kL(\tau) - rR(\tau)] d\tau \tag{4-4}$$

where $F(t)$ = fatigue level, $L(t)$ = training load, $R(t)$ = recovery time.

Tactical Adaptation: Cluster teams by playing style (e.g., offensive vs. defensive) using k-means clustering, then schedule matchups to balance competitive intensity.

(4) Validation and Extensions; Validation: Test the generalized model on real-world datasets (e.g., NBA 2023-24 schedule, FIBA World Cup 2022). Incorporate machine learning to predict team performance based on historical data. Develop a hybrid model for tournaments with both single-elimination and round-robin stages.

This modular framework enables seamless adaptation of scheduling models to diverse team sports. By parameterizing rules, temporal constraints, and team dynamics, the model supports efficient tournament design for basketball and other sports, ensuring fairness, feasibility, and maximal spectator engagement.

6 Model Strengths and Weaknesses

Model Strengths

1. The Analytic Hierarchy Process (AHP) is known for being intuitive and straightforward, making it ideal for structuring complex decision problems by incorporating both qualitative and quantitative criteria. It is particularly beneficial in group decision-making scenarios.

2. The Entropy Weight Method stands out for its objective and data-driven approach, which helps reduce subjective biases in decision-making. It is well-suited for handling complex decisions involving quantitative data.

3. The Christofides algorithm is an efficient approximation algorithm with time complexity ($O(n^3)$) and guarantees that the quality of the solution is no more than 1.5 times that of the optimal solution. Based on clear graph-theoretic steps, it is easy to implement and has a solid theoretical foundation, which makes it particularly suitable for medium-sized TSP instances that satisfy triangular inequalities.

Model Weaknesses

1. The model relies on the availability and quality of data for each variable, and incomplete or inaccurate data may affect results.

2. AHP heavily relies on expert judgments, making it subjective and potentially biased based on the opinions of decision-makers. The final results can be sensitive to the weighting process, and variations in weight assignments may lead to different outcomes.

3. For the Christofides algorithm, its approximation ratio limit of 1.5 may lead to less-than-optimal solution quality for some instances. In addition, the $O(n^3)$ time complexity is inefficient when dealing with large-scale problems, and the algorithm's dependence on graph completeness and triangular inequalities limits its applicability. In practice, the performance of the algorithm may not be as flexible and efficient as some heuristic or metaheuristic algorithms.

7 References

[1] Ströbel, Tim, Maximilian Stieler and Pascal Stegmann. "Digital transformation in sport : The disruptive potential of digitalization for sport management research." (2020).
[2] Liang, Yongxing. "The Economic, Municipal, And Social Impact of Major Sporting Events on Host Countries – The World Cup as Examples." Highlights in Business, Economics and Management (2023).
[3] Yadav, Jitendra Singh, Rambalak Yadav, Nidhi Sharma Sahore and Aparna Mendiratta. "Digital social engagements and knowledge sharing among sports fans: Role of interaction, identification, and interface." Technological Forecasting and Social Change (2023).
[4] Zimmermann, S.K., Dauter, J., & Gewald, H. (2024). Promoting Sports and Wellness with Social Media Advertising – An Analysis of Different Marketing Channels. Procedia Computer Science.
[5] Luo, L., Dai, Y., & Huang, F. (2015). Glocalization and the Rise of the Chinese Basketball Market. The International Journal of the History of Sport, 32(10), 1321–1335.
[6] Qi, Y., Sajadi, S.M., Baghaei, S., Rezaei, R., & Li, W. (2024). Digital technologies in sports: Opportunities, challenges, and strategies for safeguarding athlete wellbeing and competitive integrity in the digital era. Technology in Society.

8 Appendix

1. RRT Code (Python)

```python
def generate_initial_schedule(teams_list):
    """生成基础轮转赛程"""
    n_teams = len(teams_list)
    schedule = []

    # 固定第一个队伍为主队
    anchor = teams_list[0]
    others = teams_list[1:]

    for i in range(n_teams-1):
        for j in range(len(others)):
            home = others[j]
            away = others[(j+1) % len(others)]
            schedule.append((home, away))

        # 旋转队伍列表
        others = others[1:] + [others[0]]

    # 转换为主客场交替格式
    full_schedule = []
    for week in range(len(schedule)):
```

```
        home_teams = [match[0] for match in schedule[:week+1]] +
[anchor]
        away_teams = [match[1] for match in schedule[:week+1]]
        full_schedule.append(list(zip(home_teams, away_teams)))

    return full_schedule
```

2. Christofides Code (Python)

```
def christofides_tsp(dist_matrix):
    """Christofides 近似算法求解 TSP"""
    n = len(dist_matrix)
    G = nx.Graph()

    # 步骤 1：构建最小生成树(MST)
    mst = nx.minimum_spanning_tree(nx.from_numpy_matrix(dist_matrix))

    # 步骤 2：找到奇数度节点
    odd_degree_nodes = [node for node, degree in
dict(mst.degree()).items() if degree % 2 != 0]

    # 步骤 3：构建完全图并寻找最小权匹配
    subgraph = nx.Graph()
    for u, v in permutations(odd_degree_nodes, 2):
        subgraph.add_edge(u, v, weight=dist_matrix[u][v])
    min_matching = nx.algorithms.matching.min_weight_matching(subgraph,
maxcardinality=True)

    # 步骤 4：合并 MST 和匹配结果，生成欧拉回路
    multigraph = nx.MultiGraph(mst)
    multigraph.add_edges_from(min_matching)

    # 步骤 5：生成欧拉路径并转换为哈密顿路径
    euler_circuit = list(nx.eulerian_circuit(multigraph))
    hamilton_path = []
    for u, v in euler_circuit:
        if u not in hamilton_path:
            hamilton_path.append(u)
        if v not in hamilton_path:
            hamilton_path.append(v)

    # 转换为对称路径（往返）
    tour = []
    for i in range(len(hamilton_path)-1):
        tour.append((hamilton_path[i], hamilton_path[i+1]))
    tour.append((hamilton_path[-1], hamilton_path[0]))  # 返回起点

    return tour
```