Never Change a Winning Team – Unless It's More Fun!

Optimizing Tournament Narratives via Multi-Commodity Flow Models

Simon Stevens

SIAM Student Chapter

Agenda

1. Motivation
2. Introduction to Multi Commodity Flow Models
3. Modeling the World Cup as Multi Commodity Flow
4. Details on the FUN Parameters
5. Implementation and First Results
6. Extensions

Motivation

Why Optimize “Fun” in the FIFA World Cup?



Assume you are Gianni Infantino, FIFA president.

You are:

- Corrupt
- Greedy
- Allmighty

You want:

- MORE MONEY!!!

How do you get it?

- More sponsorships and advertising revenue
- More fans watching more matches of the tournament

$\Longrightarrow$ You want to maximize the entertainment value of the tournament.

$\Longrightarrow$ You can decide which team wins each match in the knockout stage.

World Cup Context



Target

Maximize entertainment value of the FIFA World Cup knockout stage with $|K| = 2^R$ teams.

Assumptions

- 32 teams, i.e., 5 rounds and 31 matches
- No group stage seeding or constraints $\implies$ any team can face any other in the first round

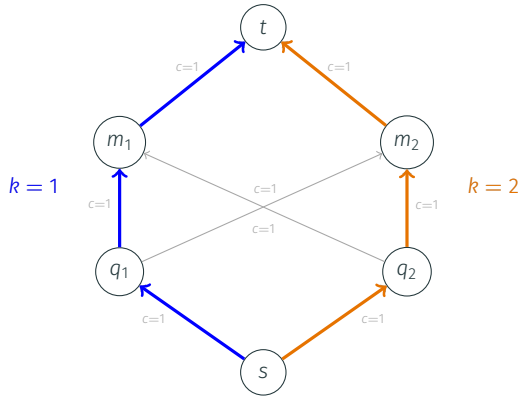
$\implies$ Tournament progression becomes a network-flow problem with one commodity per team

Introduction to Multi Commodity Flow Models

MCF Basics

- In a multi-commodity flow (MCF) model, each commodity has its own flow variables on the same network.
- Here: each team $k \in K$ is one commodity.
- Binary flow variable $x_a^k = 1$ means team k uses arc a.
- Flow conservation enforces coherent team paths from source to sink.

One Graph, Two Commodities

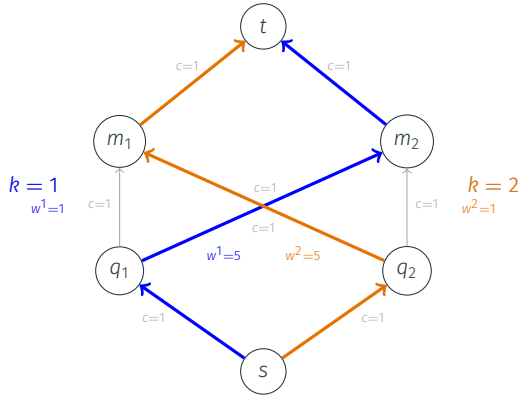


Step 1: one global source s (bottom), one global sink t (top), same graph for all commodities k.

Step 2: commodity $k = 1$ selects a feasible path to t via variables x_a^1.

Step 3: commodity $k = 2$ selects its own path via x_a^2 on the same network.

One Graph, Two Commodities



Step 4: commodity-specific weights w_a^k on the middle arcs.

Step 5: both commodities take the crossing arcs to maximise $\sum_k \sum_a w_a^k x_a^k = 5 + 5 = 10$.

General MCF Model

Given: directed graph $G = (V, A)$, commodities $k \in K$, capacities c_a, gains $w_a^k \geq 0$.

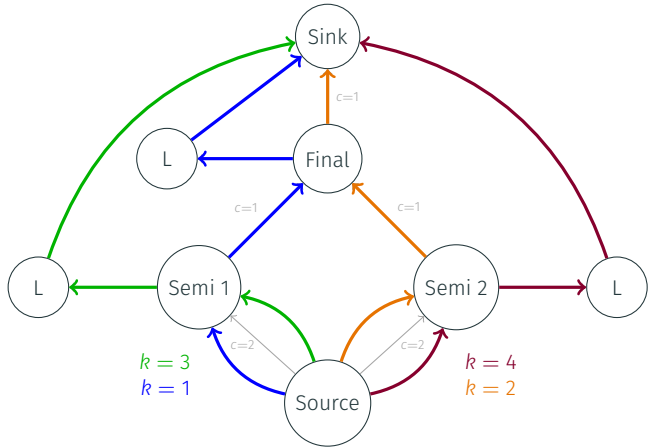
Variables: $x_a^k \in \{0, 1\}$ for all $a \in A$, $k \in K$: does commodity k use arc a?

$$
\begin{aligned}
\max \quad & \sum_{k \in K} \sum_{a \in A} w_a^k x_a^k \\
\text{s.t.} \quad & \sum_{a \in \delta^+(v)} x_a^k - \sum_{a \in \delta^-(v)} x_a^k = b_v^k && \forall\, k \in K,\ v \in V, && (1) \\
& \sum_{k \in K} x_a^k \leq c_a && \forall\, a \in A, && (2) \\
& x_a^k \geq 0 && \forall\, k \in K,\ a \in A.
\end{aligned}
$$

- (1) **flow conservation:** $b_v^k = +1$ at source, -1 at sink, 0 elsewhere.
- (2) **shared capacity:** all commodities compete for the same arc capacity.
- Integer/binary x_a^k turns this into a **MILP** — our setting.

Modeling the World Cup as Multi Commodity Flow

A Four-Team Tournament



Step 1: network structure for a four-team single-elimination tournament.
Step 2: Semifinal 1 with teams $k = 1$ (blue) vs. $k = 3$ (green); $k = 1$ wins and advances to final.
Step 3: Semifinal 2 with teams $k = 2$ (orange) vs. $k = 4$ (purple); $k = 2$ wins and advances to final.
Step 4: Final with $k = 1$ (blue) vs. $k = 2$ (orange); $k = 2$ wins the championship.
Step 5: all eliminated teams and the runner-up exit through loser nodes to the sink.

Tournament Network Construction

- Teams $k \in K$, one commodity per team.
- Nodes: source s, sink t, match nodes v_m, loser nodes ℓ_m for each match $m \in M$.
- Arc classes:
  - seeding arcs $s \to v_m$,
  - winner arcs between match nodes (and final winner to t),
  - loser arcs $v_m \to \ell_m$,
  - sink arcs $\ell_m \to t$.
- Result: a directed acyclic graph with explicit elimination structure.

Decision Variables

$x_a^k \in \{0,1\}$	team-flow on arc a,
$y_m^k \in \{0,1\}$	team k participates in match m,
$w_m^k \in \{0,1\}$	team k wins match m,
$z_m^{ij} \in \{0,1\}$	pair (i,j) meets in match m.

Objective Function

$$\max \sum_{m \in M} \sum_{k \in K} f_m^k y_m^k + \sum_{m \in M} \sum_{i<j} g_m^{ij} z_m^{ij} + \sum_{k \in K} h^k w_{m\star}^k$$

- First term rewards attractive teams surviving to later rounds (f_m^k).
- Second term rewards specific pairings (rivalries, narratives) (g_m^{ij}).
- Third term rewards a preferred champion in the final match $m^\star$ (h^k).

Variables/parameters:

- $y_m^k = 1$: team k participates in match m.
- $z_m^{ij} = 1$: teams i, j meet in match m.
- $w_m^k = 1$: team k wins match m.
- f_m^k, g_m^{ij}, h^k: fun coefficients.

Core Constraints

- **Flow conservation:** one unit per team from source to sink.
- **Participation linkage:**

$$y_m^k = \sum_{a \in \delta^-(v_m)} x_a^k$$

- **Match structure:**

$$\sum_{k \in K} y_m^k = 2, \qquad \sum_{k \in K} w_m^k = 1, \qquad w_m^k \le y_m^k$$

- **Winner/loser consistency:**

$$x_{a_m^W}^k = w_m^k, \qquad x_{a_m^L}^k = y_m^k - w_m^k$$

- **Matchup linearization:** standard three inequalities for $z_m^{ij} = y_m^i y_m^j$.

Variables/parameters:

- $x_a^k = 1$: team k flows through arc a.
- $y_m^k = 1$: team k participates in match m.
- $w_m^k = 1$: team k wins match m.
- $z_m^{ij} = 1$: teams i, j meet in match m.
- a_m^W, a_m^L: winner/loser outgoing arcs of match m.

Full Tournament MILP

$$\max \quad \sum_{m\in M}\sum_{k\in K} f_m^k\, y_m^k + \sum_{m\in M}\sum_{i<j} g_m^{ij}\, z_m^{ij} + \sum_{k\in K} h^k\, w_{m\star}^k$$

$$\text{s.t.} \quad \sum_{a\in\delta^+(v)} x_a^k - \sum_{a\in\delta^-(v)} x_a^k = b_v^k \qquad \forall\, k,\ v\in V \tag{3}$$

$$y_m^k = \textstyle\sum_{a\in\delta^-(v_m)} x_a^k \qquad \forall\, k,\ m \tag{4}$$

$$\textstyle\sum_k y_m^k = 2, \quad \sum_k w_m^k = 1, \quad w_m^k \le y_m^k \qquad \forall\, k,\ m \tag{5}$$

$$x_{a_m^W}^k = w_m^k, \quad x_{a_m^L}^k = y_m^k - w_m^k \qquad \forall\, k,\ m \tag{6}$$

$$z_m^{ij} \ge y_m^i + y_m^j - 1, \quad z_m^{ij} \le y_m^i, \quad z_m^{ij} \le y_m^j \qquad \forall\, m,\ i<j \tag{7}$$

$$x_a^k,\ y_m^k,\ w_m^k,\ z_m^{ij} \in \{0,1\} \qquad \forall\, k,\ a,\ m.$$

- (3) flow conservation per team.
- (4) participation $\leftrightarrow$ incoming flow.
- (5) two participants, one winner.
- (6) winner/loser arc consistency.
- (7) linearization of $z_m^{ij} = y_m^i y_m^j$.

Details on the FUN Parameters

Details on the FUN Parameters

Team Fun Parameter f_m^k
The better / more popular a team, the higher the base fun score.

Matchup Fun Parameter g_m^{ij}
Default fun scores for all matchups, with curated boosts for specific rivalries.

Winner Fun Parameter h^k
Bonus for a preferred champion winning the final match.

Round Multipliers r_m
Later rounds get amplified fun scores to reflect higher stakes and visibility.
(Contained in f_m^k and g_m^{ij} in our model.)

$\Longrightarrow$ More interesting fun functions next week! ;)

Implementation and First Results

Model Size for 32 Teams

For $|K| = 32$, $R = 5$, $|M| = 31$, and $|A| = 125$:

- x variables: $|K| \cdot |A| = 32 \cdot 125 = 4000$.
- y variables: $|M| \cdot |K| = 31 \cdot 32 = 992$.
- w variables: $|M| \cdot |K| = 31 \cdot 32 = 992$.
- Implemented z variables: $|M| \cdot |K|^2 = 31 \cdot 32^2 = 31\,744$.

$\Longrightarrow$ Overall binary variables: $4000 + 992 + 992 + 31\,744 = 37\,728$.

How to Solve the Model Faster

- **Cut redundant z variables early:** create z_m^{ij} only for feasible team pairs at match m.
- **Add symmetry-breaking constraints:** enforce canonical ordering rules in early rounds to reduce equivalent brackets.
- **Warm-start with a heuristic bracket:** seed Gurobi with a heuristic solution based on fun scores.
- **Decompose by rounds/subtrees:** solve regional subproblems first, then coordinate at upper-round matches.

For 5 rounds, we do not need these techniques, BUT:
$\Rightarrow$ For $R = 10$ ($|K| = 1024$, $|M| = 1023$, $|A| = 4093$): $x = 4\,191\,232$, $y = w = 1\,047\,552$, $z = 1\,072\,693\,248$, total $= 1\,078\,979\,584$ binaries.

Observed Qualitative Behavior

- High-base teams are pushed deeper because round multipliers are increasing.
- Rivalries are prioritized when bracket feasibility allows their meeting.
- Winner bonus biases the final champion toward globally attractive teams.

Extensions

Possible Research Extensions

- Add fairness constraints (e.g., confederation separation, rematch restrictions).
- Add constraints to model group stage seeding and its impact on knockout fun.
- Incorporate probabilistic match outcomes based on data-driven team strength models.
- Incorporate more realistic data-based fun functions.

  $\Longrightarrow$ Increase realism by only being able to influence group stage seedings.
- Use robust optimization for uncertain fun scores and audience preferences.
- Couple with TV scheduling and time-zone aware audience models.

Next week: more complex fun functions and model extensions!

“I spent a lot of money on booze, birds and fast cars. The rest I just squandered … on symmetry-breaking constraints.”

— George Best, football legend
and symmetry-breaking enthusiast

