Lineup Cohesion Score

A Network-Based Metric for Optimal Starting Eleven Selection

NEU Sports Analytics Hackathon

Prompt A: Starting Eleven Lineup Construction

February 2026

Data: IMPECT Open Data • Bundesliga 2023/24 • 306 Matches

The Challenge: Optimal Lineup Selection

How do coaches select a starting eleven that maximizes team performance?

Key Questions

- Which players connect well together?
- Who are the critical hub players?
- What passing patterns lead to goals?

Our Approach

- Build player-to-player pass networks
- Weight edges by shot creation
- Quantify lineup "cohesion"

💡 Key Insight

Elite teams are HUB-DEPENDENT: they funnel play through star players

(Counter to the intuition that "balanced" teams perform better)

Methodology: Lineup Cohesion Score

$$\text{Cohesion} = 0.50 \cdot \text{Connectivity} + 0.25 \cdot \text{Chemistry} + 0.15 \cdot \text{HubDep} + 0.10 \cdot \text{Progression}$$

Connectivity
(50%)

Network density +
avg clustering coefficient

How interconnected are the players?

Chemistry
(25%)

Critical position pair
pass frequency

Midfield→Wing, Midfield→Striker links

Hub Dependence
(15%)

Gini coefficient of
degree distribution

Star player reliance (GOOD)

Progression
(10%)

Pre-shot pass ratio
(passes → shots)

Attacking effectiveness

$$\text{Edge Weight} = \text{PassCount} \times (1 + \text{PreShotRatio})$$

Pipeline:

`IMPECT Events → Filter Passes → Build DiGraph → Compute Metrics → Aggregate Season`

* Weights optimized empirically: original "Balance" inverted to "Hub Dependence" based on correlation analysis showing elite teams are MORE centralized, not less.

Validation: Cohesion Predicts Season Performance

✓ Season-Level

r = 0.728, p < 0.001

Explains 53% of variance
in season points

✓ Match-Level

F = 36.64, p < 0.0001

Significant difference
win vs draw vs loss

Component Correlations with Season Points

Component	r	p-value	Sig
Connectivity	+0.785	0.0001	***
Hub Dependence	+0.714	0.0009	***
Chemistry	+0.448	0.0623	*
Progression	+0.133	0.5978	

Key Finding: Connectivity is the strongest predictor

Dense, well-clustered passing networks strongly associated with success

The "Balance Paradox": Star Players Matter

Original Assumption

"Balanced" teams perform better

(Even pass distribution)

r = -0.714 with points

→

Empirical Reality

Hub-dependent teams win more

(Star player centralization)

r = +0.714 with points ✓

Evidence from Top Teams:

Leverkusen (90 pts)	Xhaka orchestrates; Wirtz finishes
Bayern (72 pts)	Kimmich as central hub
Stuttgart (75 pts)	Clear passing hierarchy

Implication for Lineup Selection

Optimize for HUB CONNECTIVITY, not equal distribution.

Build lineups around your best playmaker.

Case Study: Leverkusen's Undefeated Season

2023/24 Season

28W - 6D - 0L

90 points | +63 GD

87 goals scored

Cohesion Score

0.560

League avg: 0.548

Hub Players

Granit Xhaka

(Volume hub: 558 passes)

⚡ Florian Wirtz

Top Attacking Connections (Pre-Shot Passes)

Connection	Pre-Shot	Conversion
Wirtz → Boniface	**22**	23.4%
Frimpong → Boniface	**12**	38.7%
Xhaka → Wirtz	**11**	6.1%
Palacios → Boniface	**9**	15.5%

Tactical Pattern: The Leverkusen Chain

`Defense → Xhaka (orchestrate) → Wirtz (create) → Boniface (finish)`

Wirtz leads ALL pre-shot connections — the attacking fulcrum

Application: Lineup Optimization Framework

For Coaches

- Identify optimal player combinations
- Quantify impact of substitutions
- Plan for injuries: who maintains hub connectivity?

For Analysts

- Compare lineup alternatives
- Scout opposition networks
- Evaluate transfer targets' fit with existing passing structure

Example: Evaluating a Lineup Change

Lineup A (Current)

Cohesion: 0.52

Connectivity: 0.31

→

Lineup B (Proposed)

Cohesion: 0.58 (+12%)

Connectivity: 0.36

⚠ Limitations

• Requires historical pass data • Does not capture off-ball movement

• Single season validation • Opponent-specific effects not modeled

Conclusions & Future Work

Key Contributions

1. Validated Cohesion Metric	r = 0.728 with season points ($p < 0.001$)
2. Hub Dependence Insight	Elite teams centralize play through stars
3. Pre-Shot Weighting	Edges weighted by attacking contribution
4. Leverkusen Case Study	Xhaka→Wirtz→Boniface chain identified

Future Directions

- Incorporate opponent-specific adjustments
- Add temporal dynamics (fatigue, momentum)
- Extend to multi-season validation
- Build interactive lineup optimizer tool

Thank You!

Code: github.com/[your-repo] | Questions?