



FIFA WORLD CUP  
Qatar2022



# ARGENTINA

**How did they win the World Cup final?  
A network-based approach**

Davide Cacciatore - 2015641  
Andrea Potì - 2008416

Data Driven Modelling of Complex Systems  
Sapienza University of Rome



SAPIENZA  
UNIVERSITÀ DI ROMA

# Table of contents

- 01. Introduction**
- 02. Related works and data extraction**
- 03. Passing networks**
- 04. Player metrics**
- 05. Zone-Player network**
- 06. Community detection**
- 07. Zone network**
- 08. Conclusions**



# Introduction

By analyzing the Argentine team's network of passes during the final match, we aim to shed light on its tactics and demonstrate the complexity of its game.



We want to challenge the media's claims that Argentina's path to victory was purely based on counterattacks.

By doing so, we hope to provide a more nuanced understanding of Argentina's triumph and contribute to the conversation about the use of network analysis in sports research.

- Argentina-Francia, la garra contro la qualità, *Corriere.it*: <https://tinyurl.com/4hscydwx>
- Messi dipinge calcio: l'Argentina va in finale nei mondiali del contropiede, *Calciomercato.com*: <https://tinyurl.com/2s3v9xda>
- Il possesso palla fa ridere. Questo Mondiale è solo difesa e contropiede, *Calciomercato.com*: <https://tinyurl.com/mvp2y45u>

# Related works and data extraction

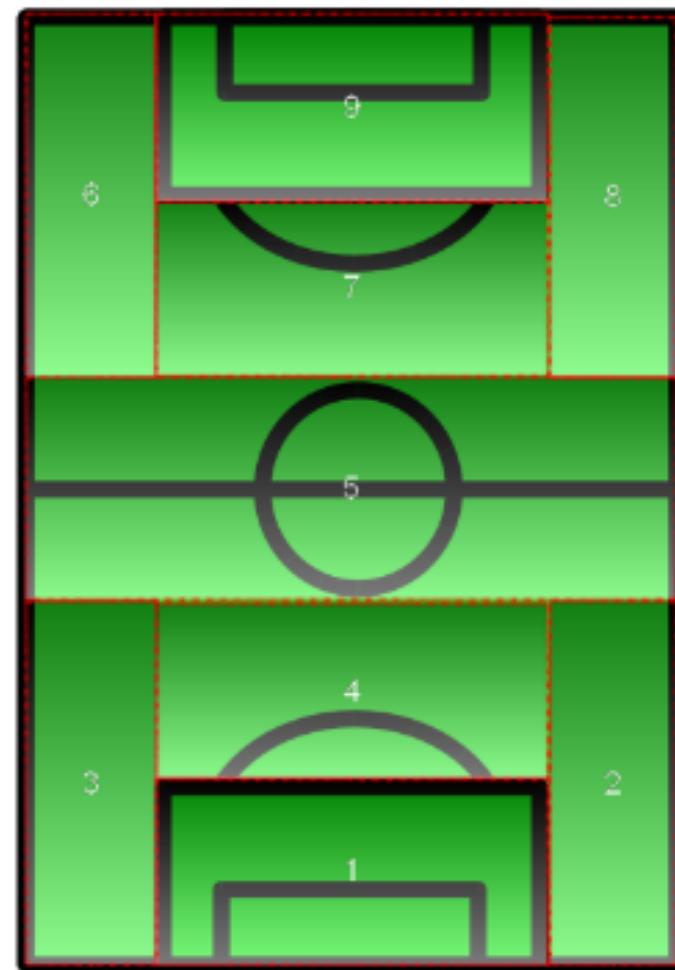


Figure 1: Model of the football field for data extraction.  
Playing zones.

Several papers try to apply complex network analysis to the football world. In particular, we followed the work done by *Cotta et al. (2011)* on the Spanish team during the 2010 World Cup. We used their 9 zones division of the pitch and their model to collect the data manually, then we replicate some of their analysis.

Then, we followed some works done by *Arriaza-Ardiles et al. (2018)* for the interpretation of the player metrics results and by *Cintia et al. (2015)* for the mean node degree as evaluation of the performance of the team.

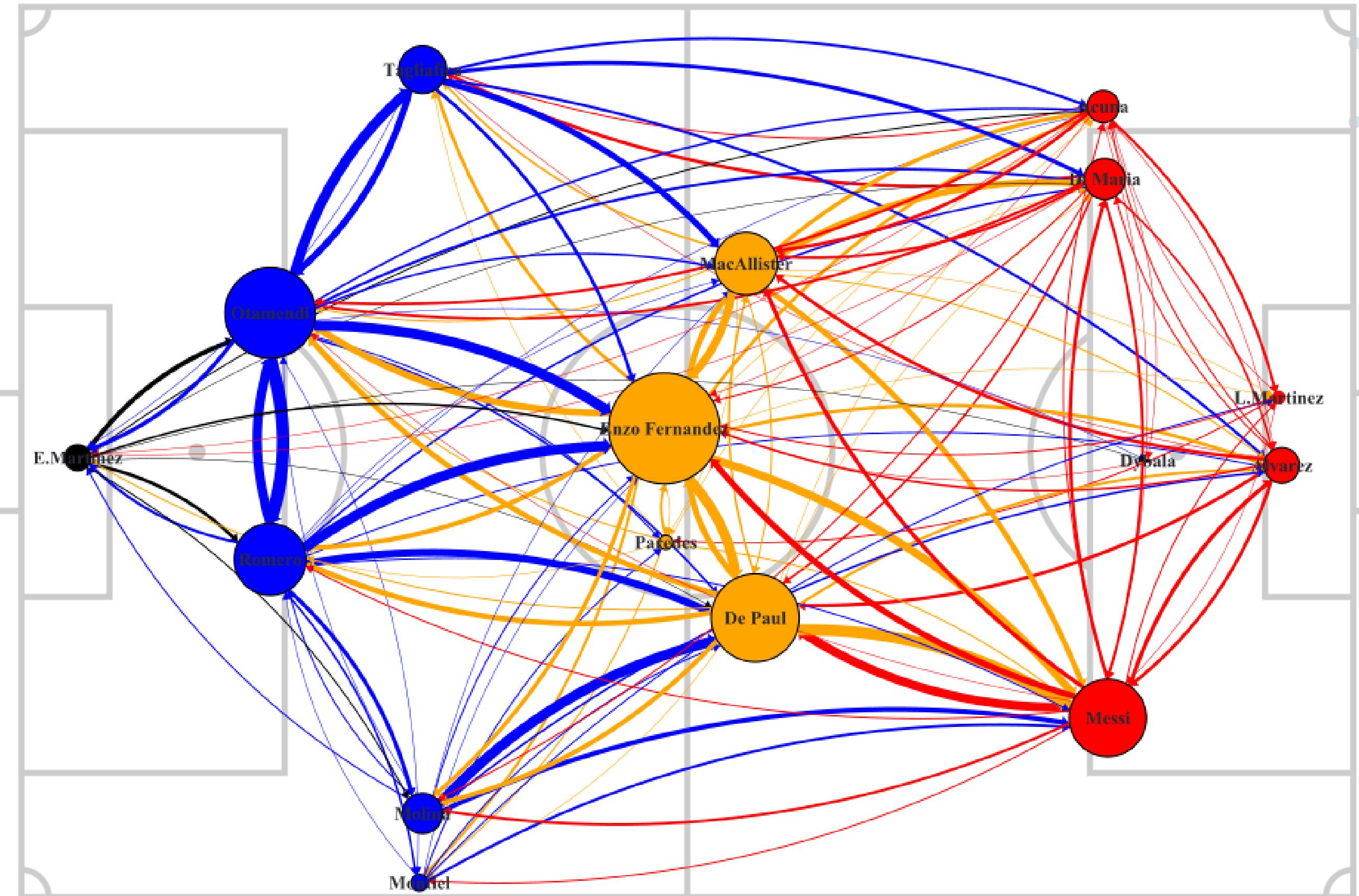
We collected the data manually by watching the entire match and they consist of a list of passes with the format:

$\langle \text{player1} \rangle \langle \text{player2} \rangle \langle \text{zone1} \rangle \langle \text{zone2} \rangle \langle \text{minute} \rangle$

indicating which player passed to which player, in which minute, and in which pitch zones were each of them.

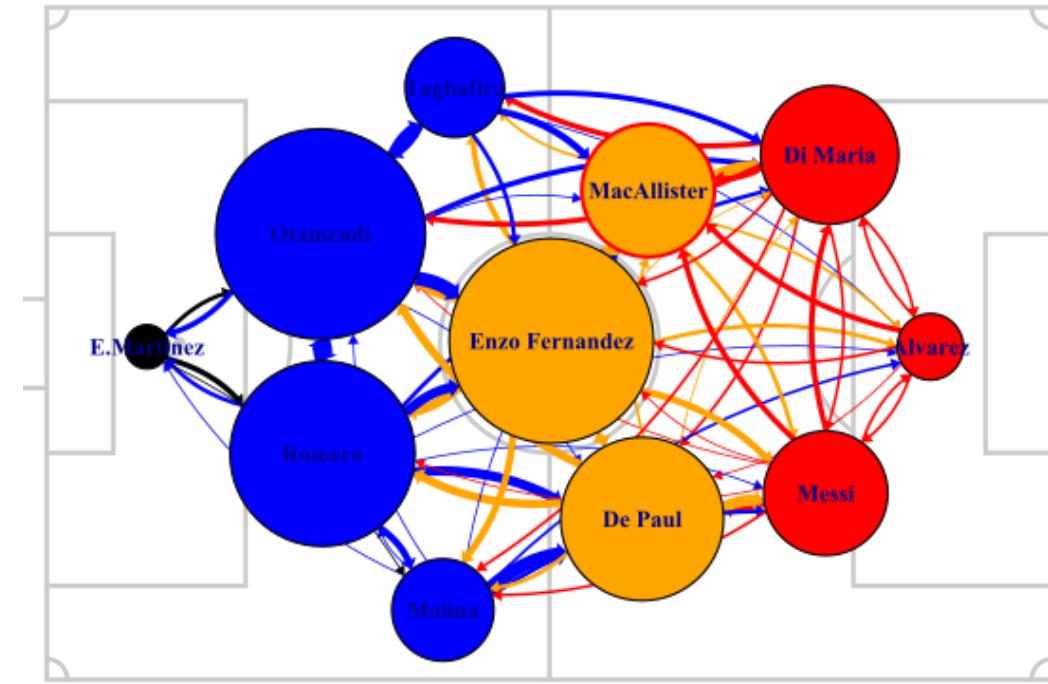
- Arriaza-Ardiles E., Martín-González J.M., Zuniga M.D., Sánchez-Flores J., de Saa Y., García-Manso J.M., Applying graphs and complex networks to football metric interpretation, *Human Movement Science*, Volume 57, 2018, Pages 236–243, ISSN 0167-9457.
- Cintia, Paolo & Rinzivillo, Salvatore & Pappalardo, Luca. (2015). A network-based approach to evaluate the performance of football teams.
- Cotta, Carlos & Mora, Antonio & Molina, Cecilia & Merelo Guervós, Juan. (2011). FIFA World Cup 2010: A Network Analysis of the Champion Team Play. *CoRR*.

# Passing network

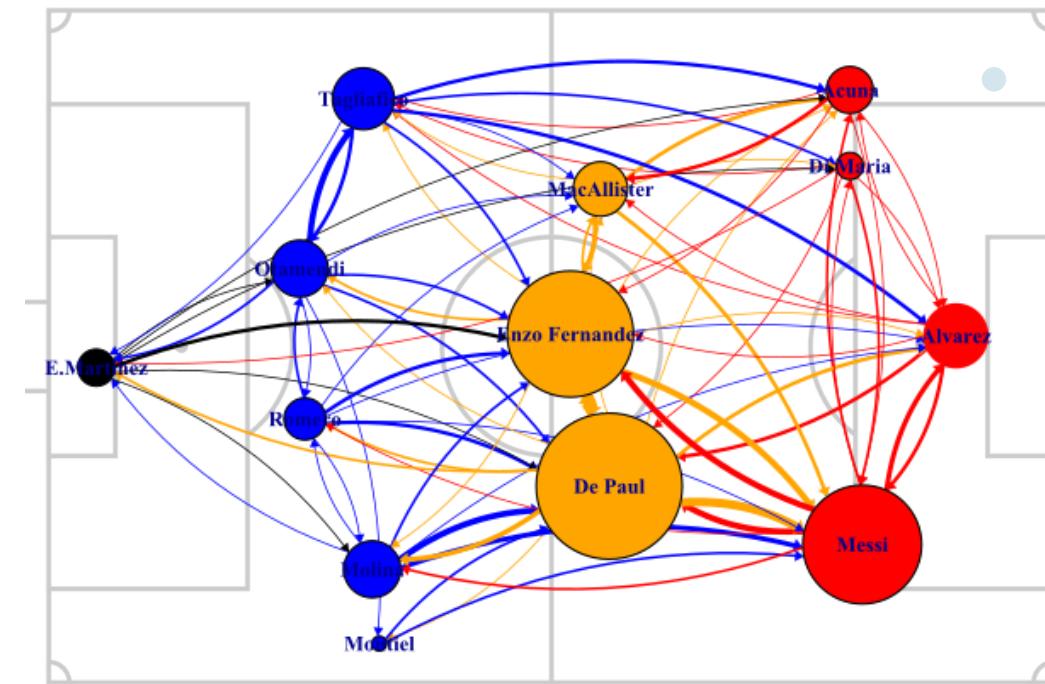


# Passing networks

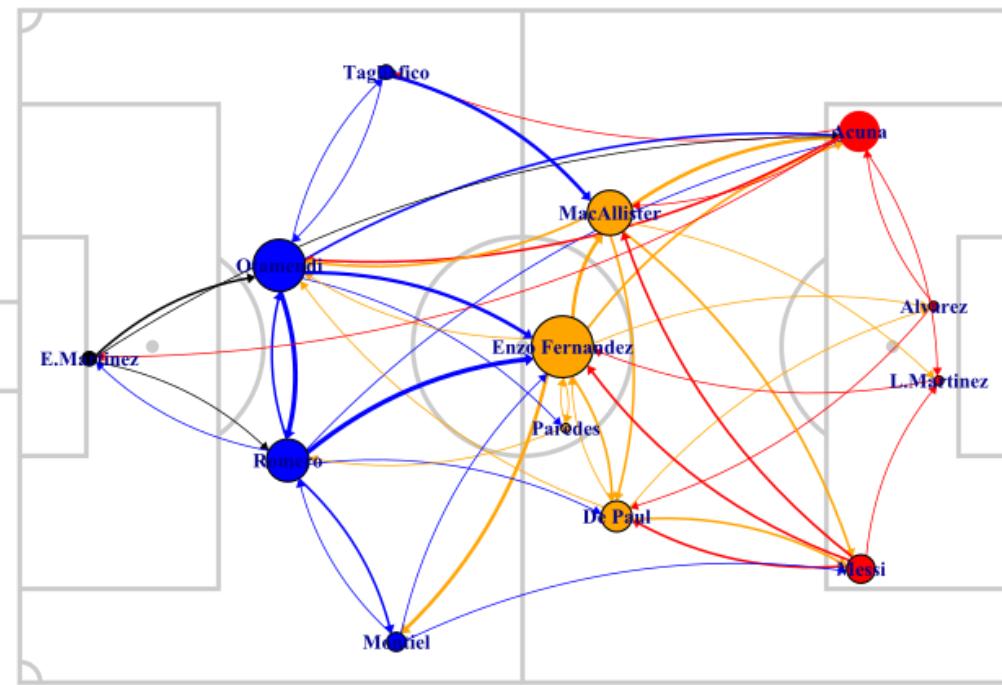
**1st half network**



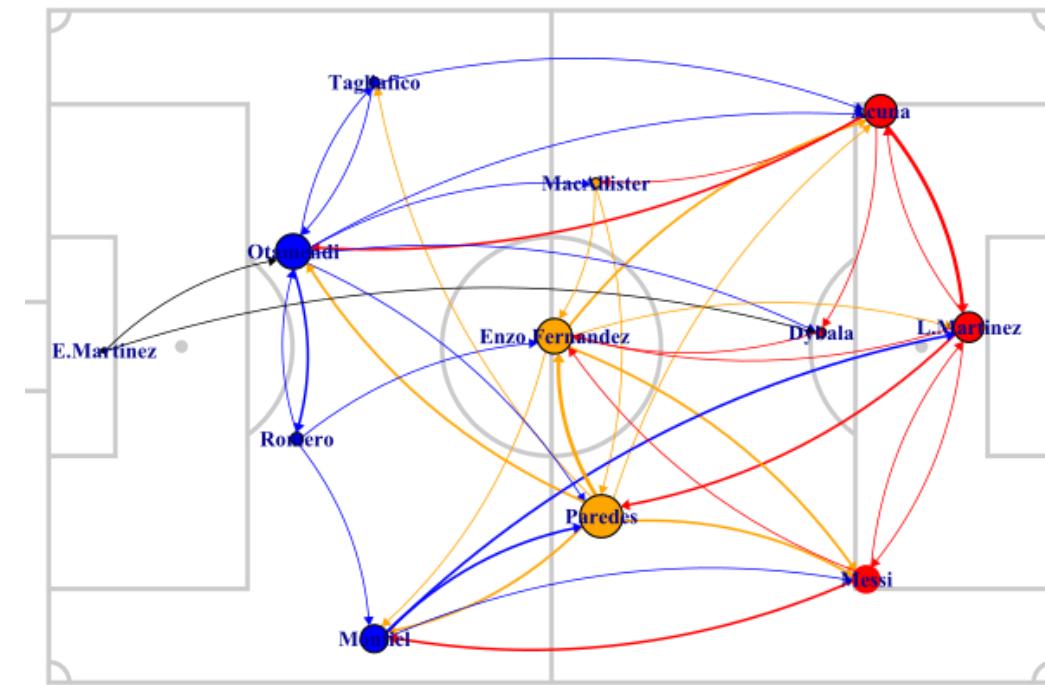
**2nd half network**



**extra-time 1st half network**



**extra-time 2nd half network**



# Player metrics

## Some considerations

Analyzing metrics gives us the possibility to explore:

- Team characteristics;
- Type of players;
- Play scheme and style of play;

The **centrality measures** give informations about the players who act as attractors of the play:

- **Closeness** regarding the contribution in terms of links generated with other players. A player has a higher centrality when he depends less from the rest to communicate with other components of the team.
- **Betweenness** in terms of intermediation between two other players. A player with high betweenness has a relevant influence during possession acting as a bridge between other players.

The **clustering coefficient** accounts for the transitivity of the network by counting all the possible triangulations related with the player. This measure allow us to discriminate the connection level of the players when the team is in possession of the ball.



# Player metrics



**Enzo Fernandez:** the key man in Argentina's game.



	Clustering coefficient	Betweenness centrality	Closeness centrality - in	Closeness centrality - out
Molina	0.111	0.003	0.625	0.682
De Paul	0.218	0.035	0.750	0.789
Romero	0.258	0.021	0.682	0.789
Otamendi	0.346	0.094	0.789	0.789
Tagliafico	0.250	0.007	0.652	0.652
Di Maria	0.178	0.009	0.714	0.714
Mac Allister	0.273	0.025	0.750	0.750
Messi	0.273	0.061	0.789	0.789
Enzo Fernandez	0.390	0.191	1.000	0.882
Alvarez	0.167	0.004	0.714	0.652
E.Martinez	0.278	0.009	0.625	0.682
Acuna	0.364	0.046	0.789	0.714
Montiel	0.190	0.005	0.625	0.600
Paredes	0.222	0.006	0.600	0.652
L.Martinez	0.133	0.003	0.600	0.577
Dybala	0.000	0.000	0.536	0.500

# Zone-Player network

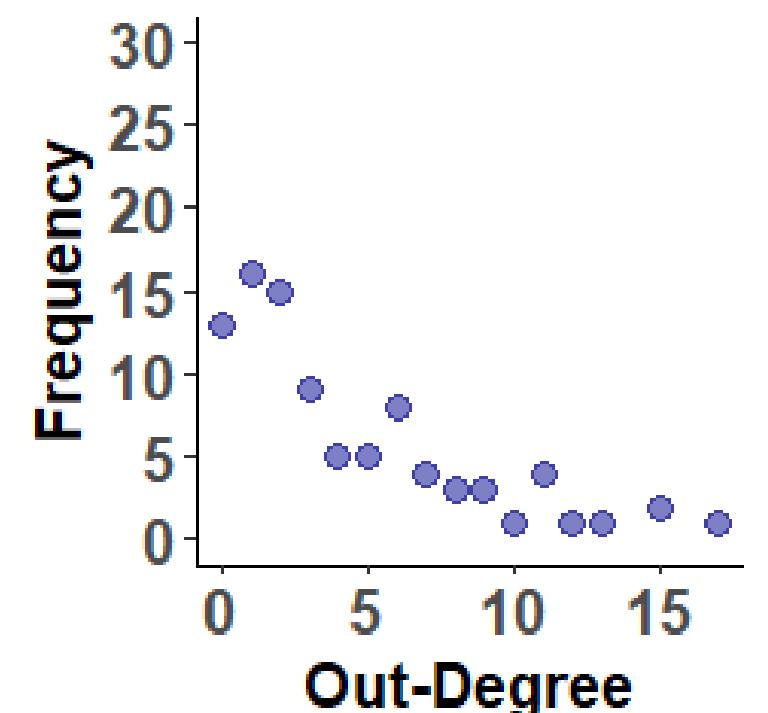
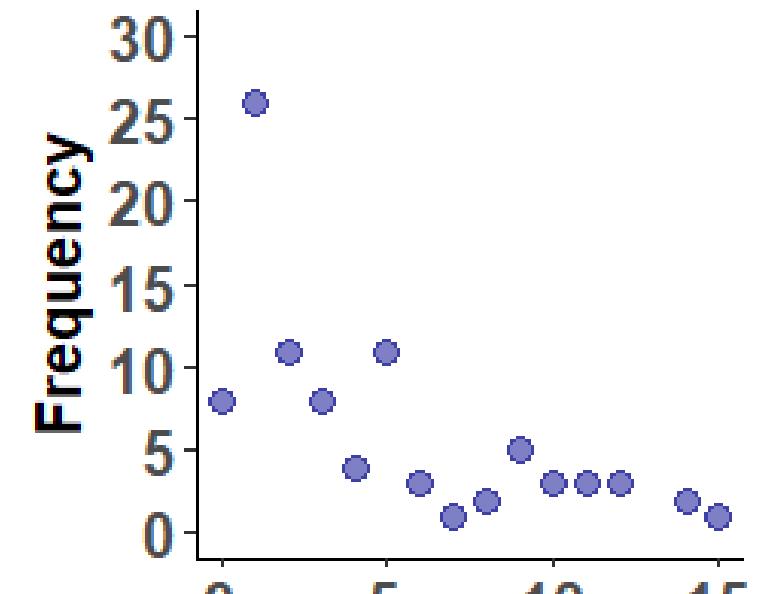
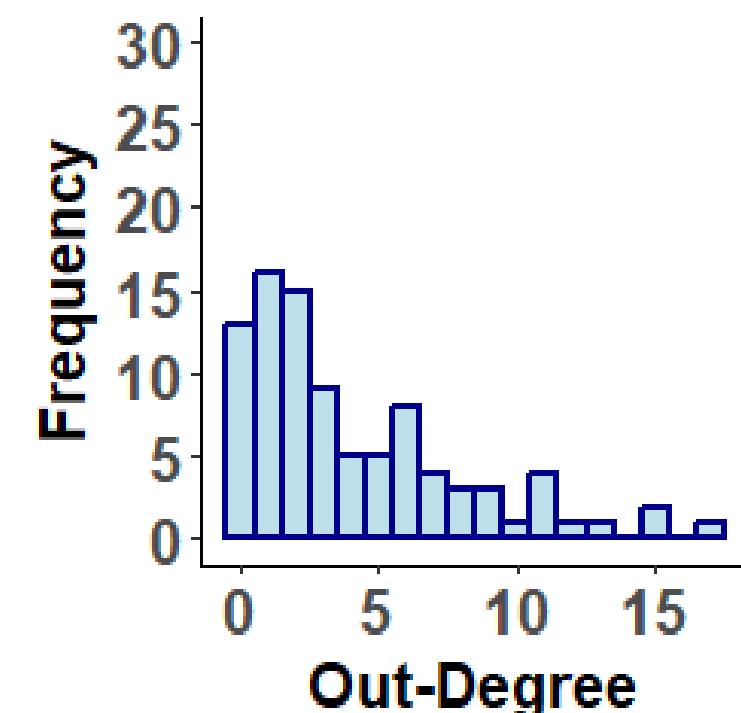
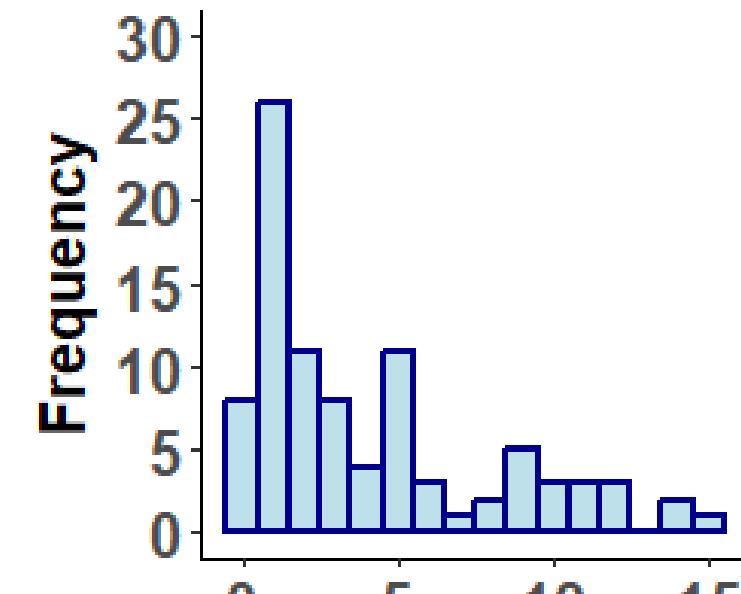
Each pair (**player, zone**) has been interpreted as a virtual player for the purposes of network construction.

This kind of analysis tries to capture the fact that a player can act in different parts of the pitch and his role may be different in each of these zones.

We built a directed graph  $G(V,E)$  where  $V=\{(p,z) / p \in T, z \in \{1,\dots,9\}\}$ ,  $T$  is the set of players in the national team and  $(u \rightarrow v) \in E$  if player  $u$  passed to player  $v$ .

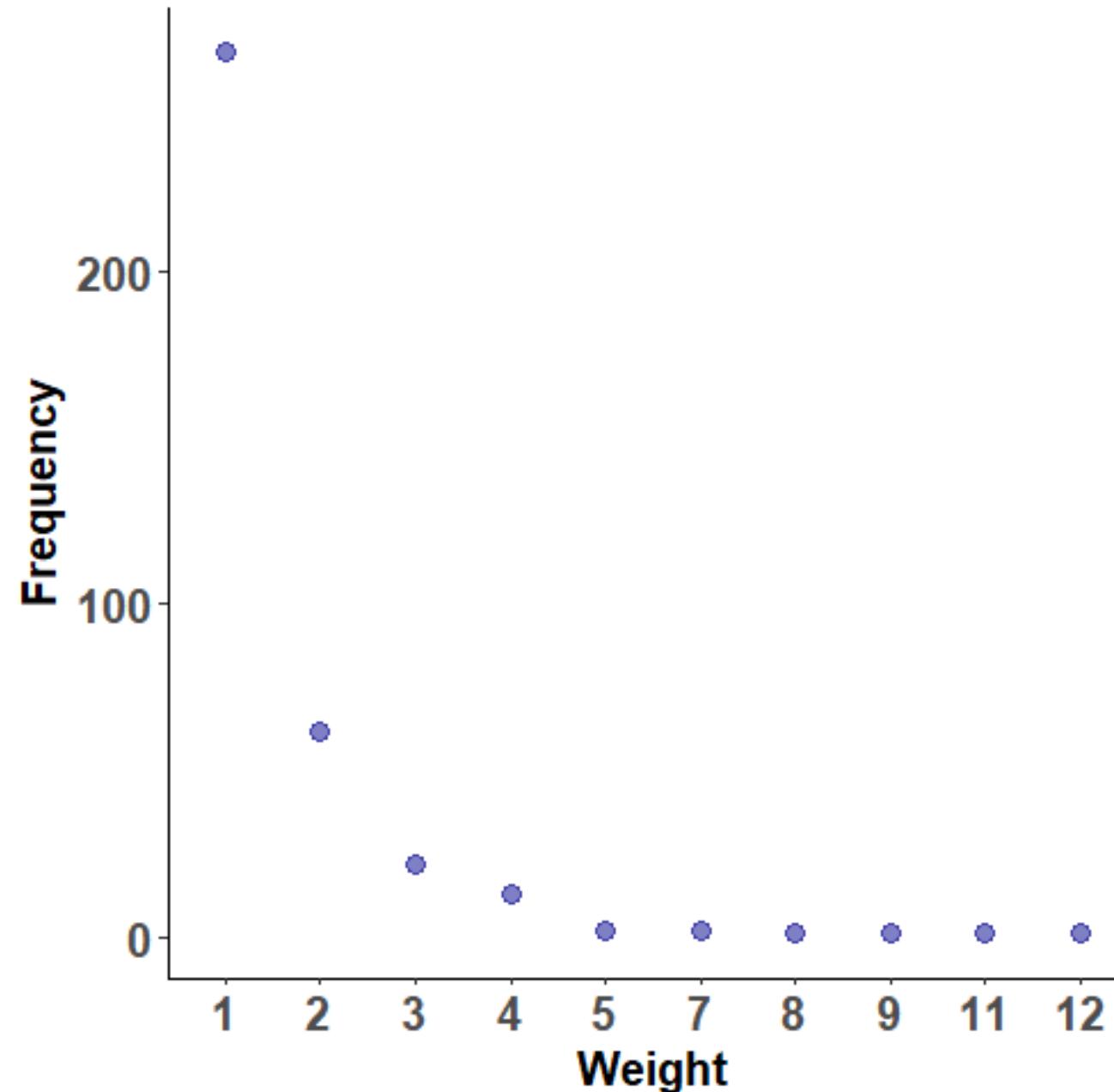
The **weight** of each edge is given by the number of passes between the two players.

## In-Degree & Out-Degree distribution



# Zone-Player network

## Strength distribution



Examining the strength distribution, we can observe that:

- Edges' weights range from 1 to 12;
- Over 200 edges are only repeated once;
- Nodes representing **Otamendi** and **Romero** in zone 4 have the strongest edges.

For the next analysis, we considered a moving window of 15 minutes to compute some metrics of the network.



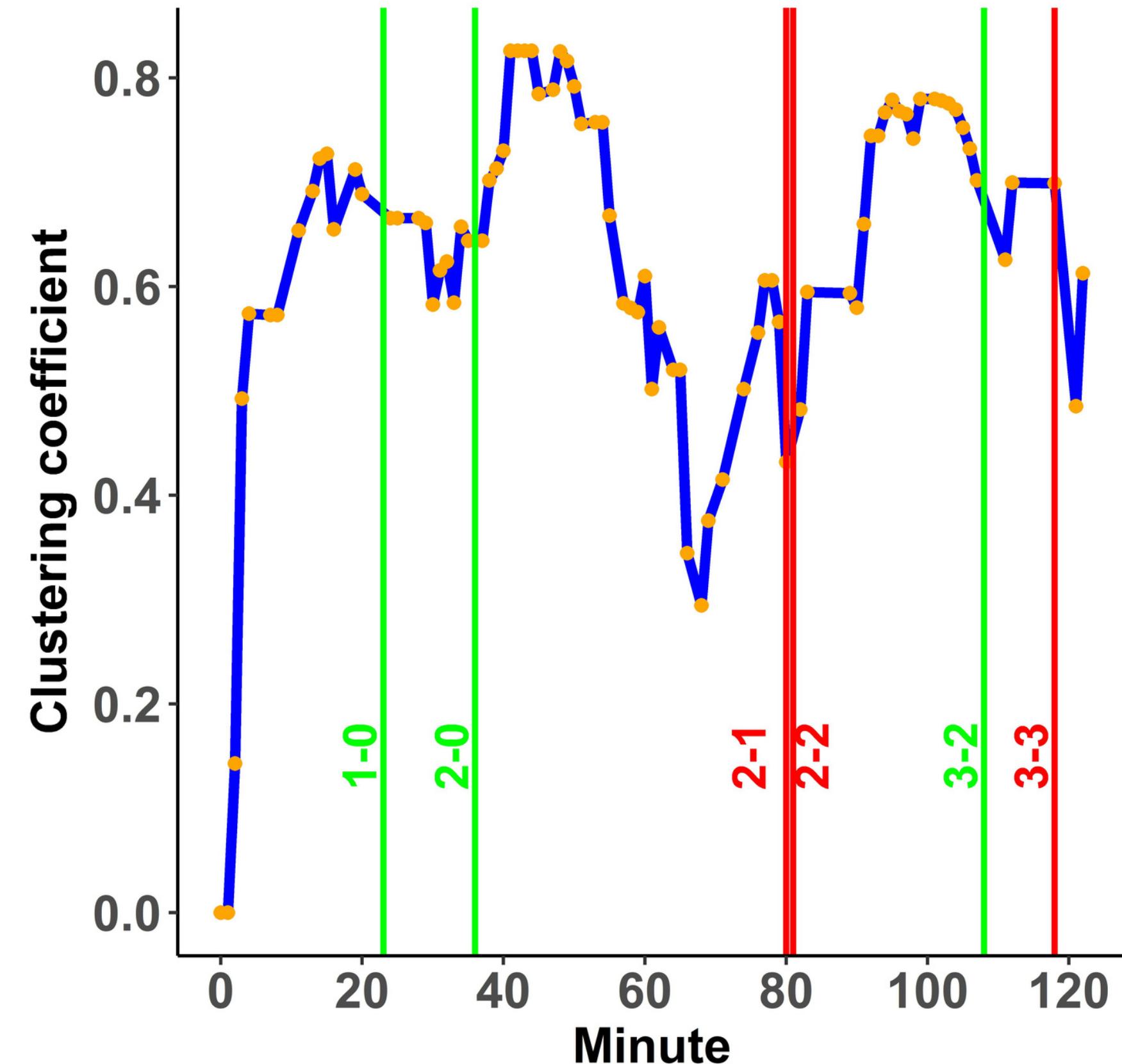
# Zone-Player network

15 minutes moving window: **Clustering coefficients**

- X-axes: Minute of the match going from 0 to 124
- Y-axes: Clustering coefficient
- **Green lines**: Argentina goals
- **Red lines**: France goals

The **Clustering coefficient** is a measure of the degree to which nodes in a graph tend to cluster together, in particular here we used the *global* type.

This metric is useful to have an idea of the ball possession level of the team, which assumes low values immediately before France goals.



# Zone-Player network

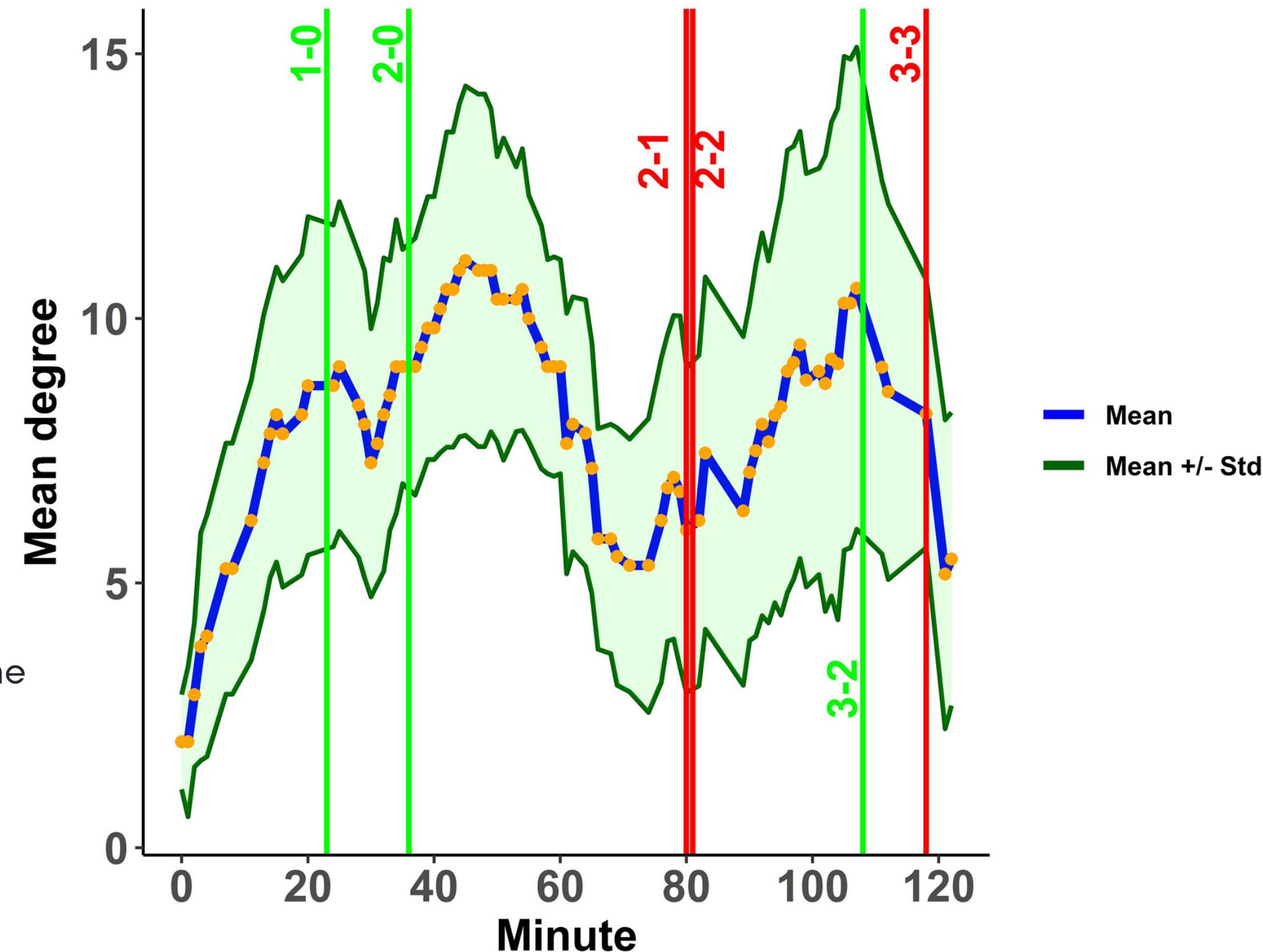
15 minutes moving window: **Mean node degree**

- X-axes: Minute of the match going from 0 to 124
- Y-axes: Mean node degree
- **Green lines**: Argentina goals
- **Red lines**: France goals

The blue line follows the **mean degree** of the players (nodes) during the match, represented by the orange points. In this case we considered the network as undirected.

We have also highlighted in green the area between one standard deviation above and below the mean.

Also in this case the value of the mean degree drops immediately before the France goals.



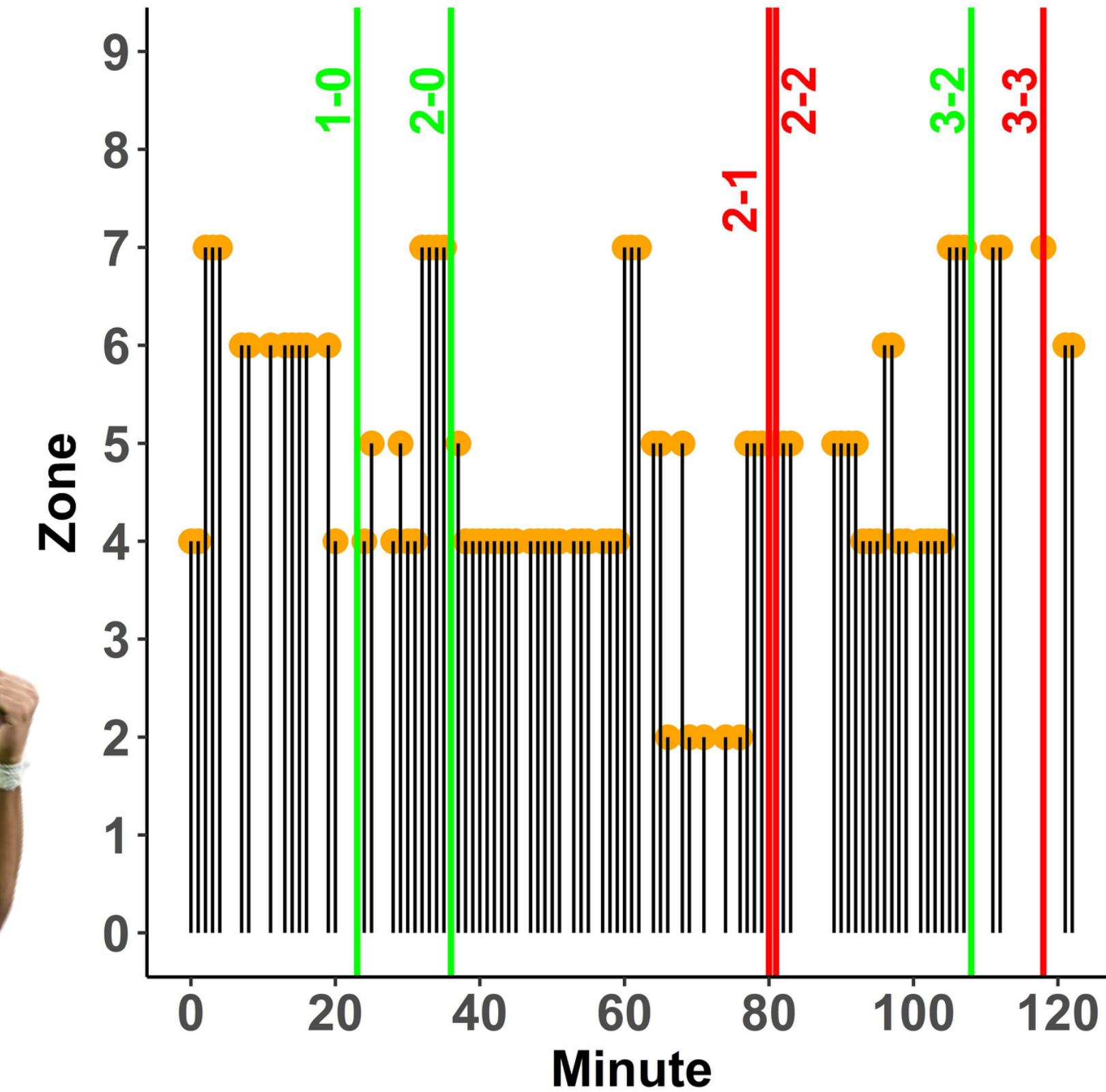
# Zone-Player network

15 minutes moving window: **Zone of the player with highest degree**

- X-axes: Minute of the match going from 0 to 124
- Y-axes: Zone of the field (low numbers zones are defensive areas, instead higher zones are the more offensive ones)
- **Green lines**: Argentina goals
- **Red lines**: France goals

Argentina's center of play is very high in the minutes corresponding to the goals.

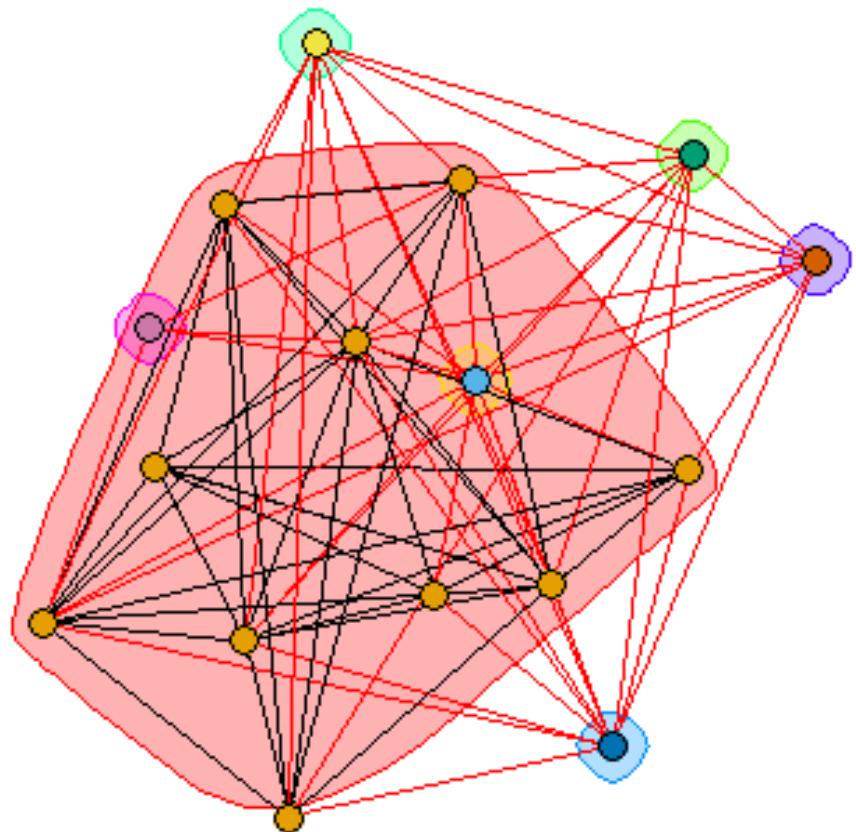
Around 60'-80', before France goals, the zone value falls to 2, which is the zone corresponding to the right back Molina, responsible for defending the attacks of **Mbappè**, the French talented left wing who scored two goals in two minutes immediately after the 80'.



# Community detection

On the player's network

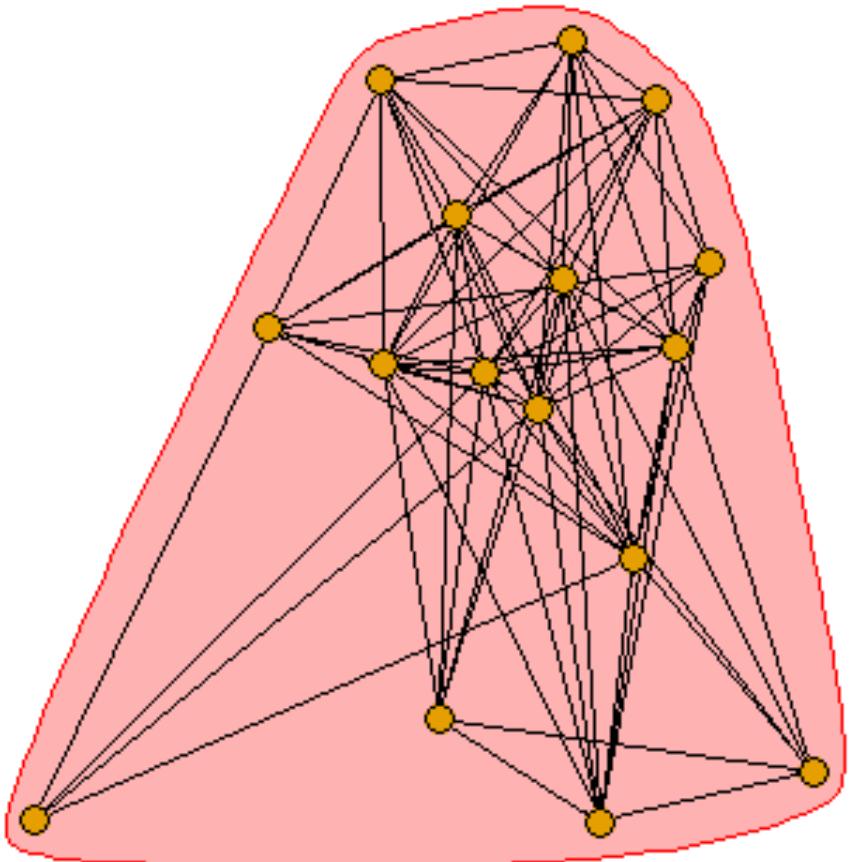
**Edge-betweenness**



7 Communities

Modularity = 0.03

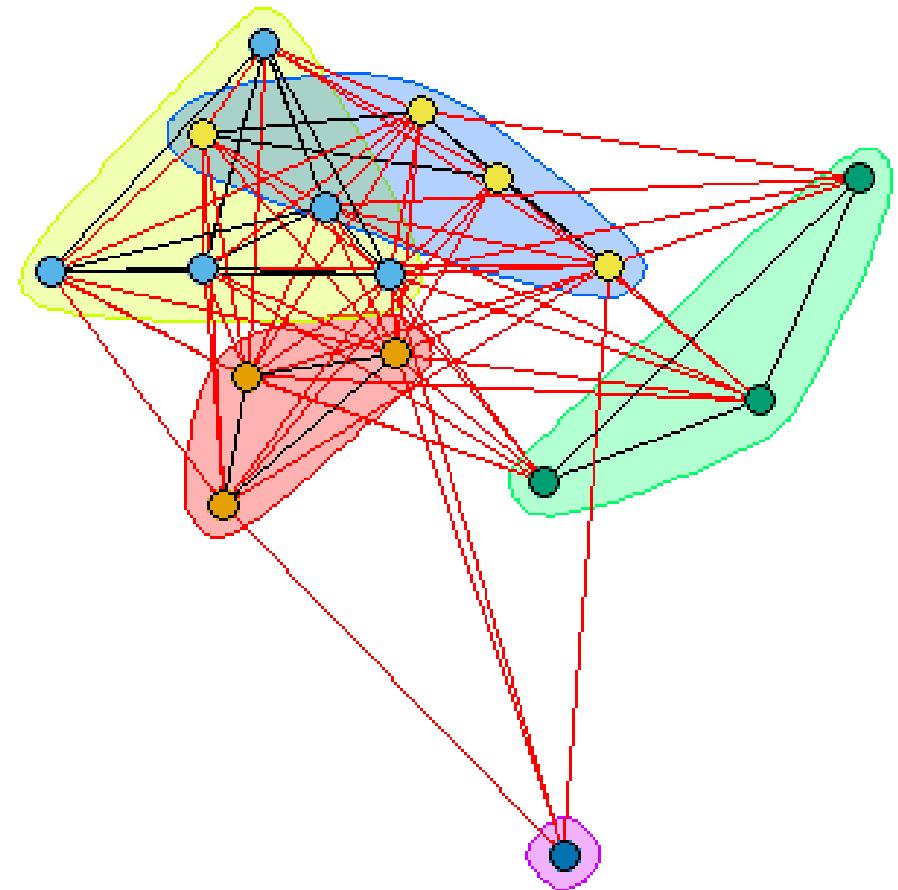
**Label-propagating**



1 Community

Modularity = 0

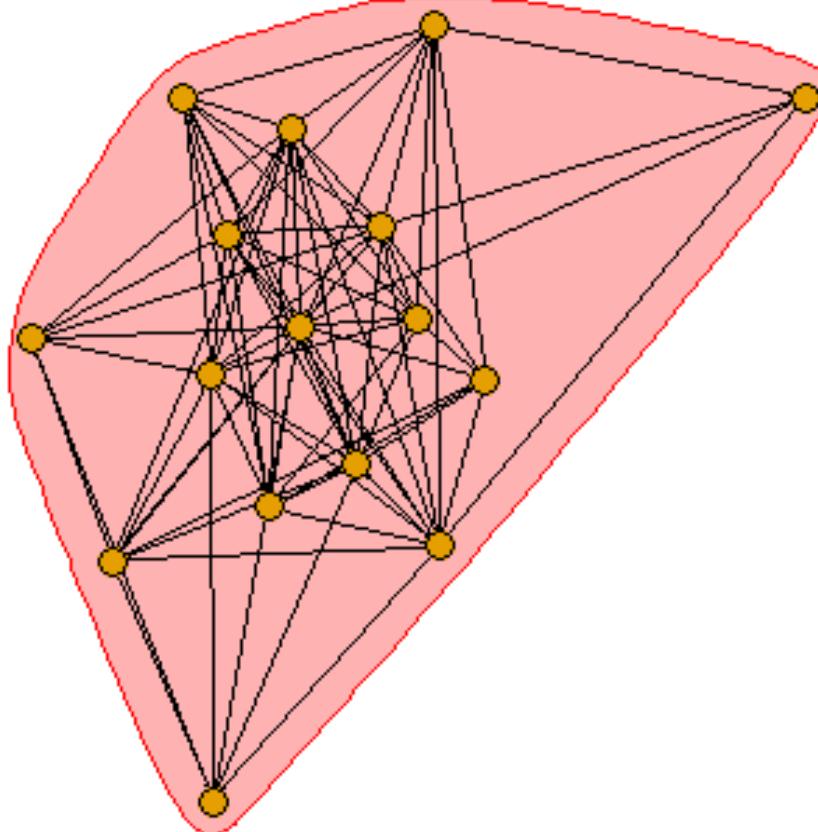
**Walktrap**



5 Communities

Modularity = 0.13

**Infomap**



1 Community

Modularity = 0

# Community detection

## Interesting results

### Infomap and Label-propagating

- Same result of one big community.

### Edge-betweenness

- Communities different from the biggest one are composed by players alone.

Acuna, Di Maria, Mac Allister, Tagliafico, Alvarez, Dybala

### Walktrap

- Communities related to play position on the field.

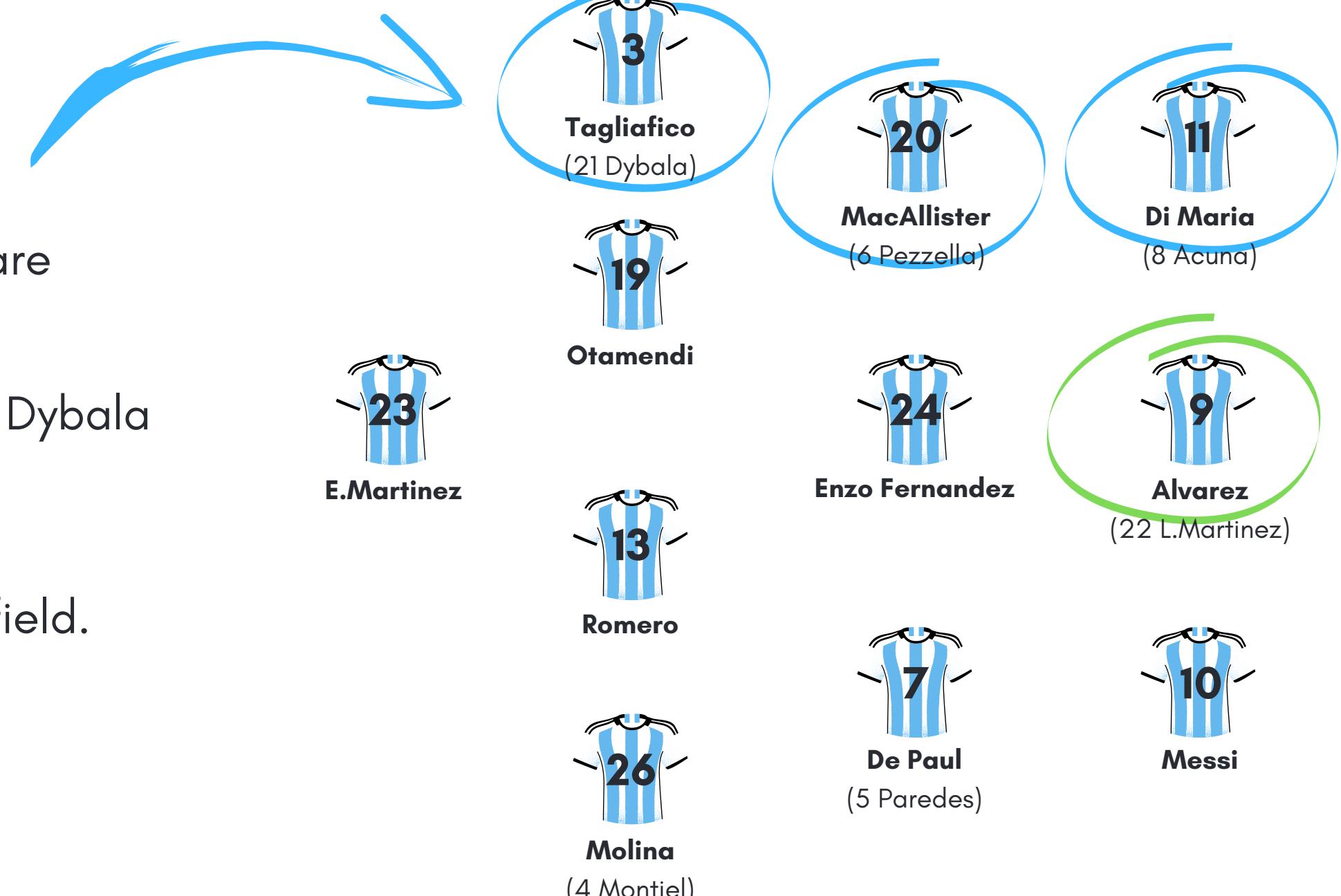
E.Martinez, Otamendi, Romero,

Enzo Fernandez, De Paul, Messi, Molina, Alvarez,

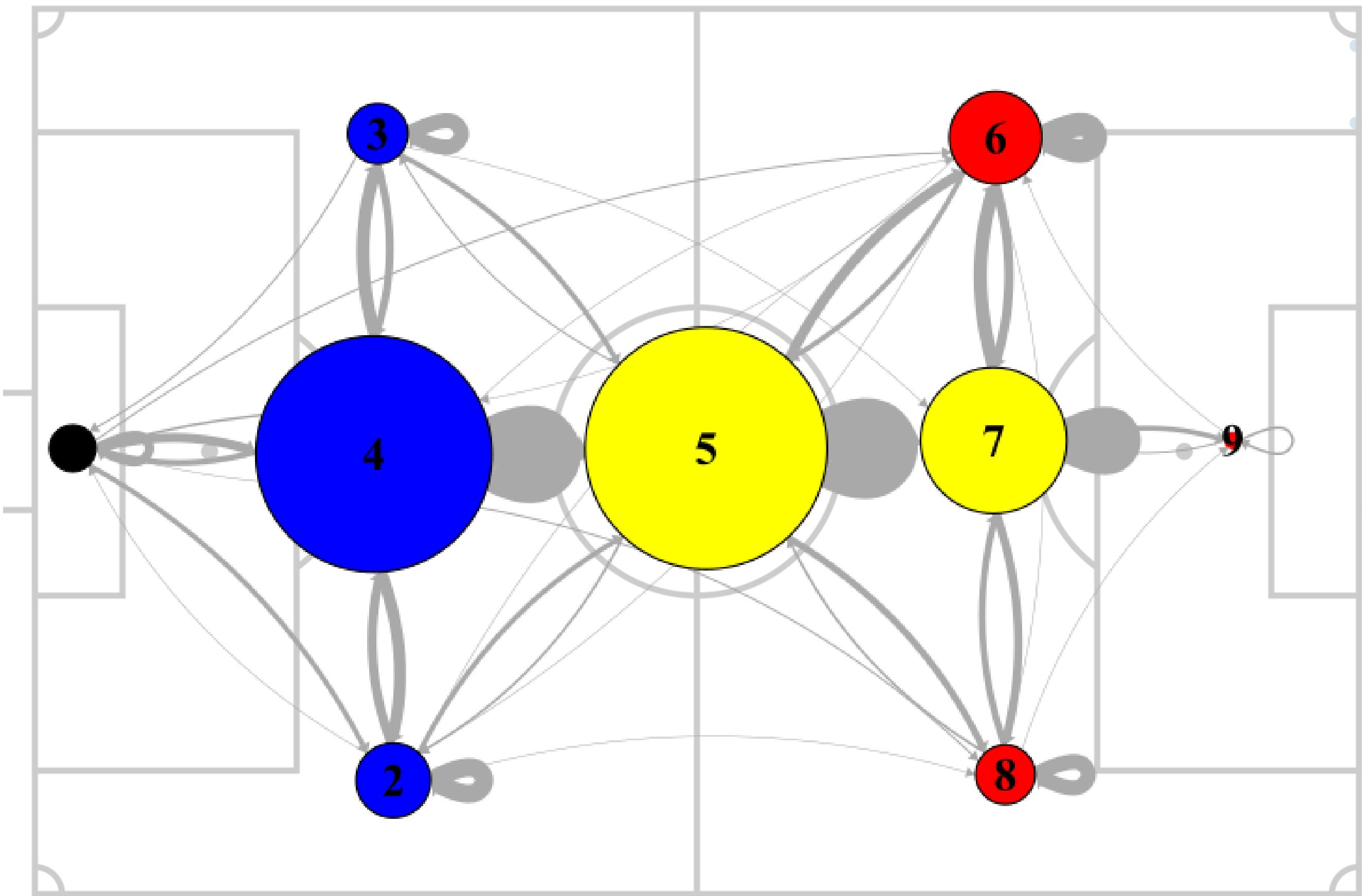
Paredes, Montiel, L. Martinez,

Acuna, Di Maria, Mac Allister, Tagliafico,

Dybala



# Zone network



# Conclusions

- We can say that the Argentina national team made its best performance in the final match of the World Cup, which they won with a deserved victory over France.
- The match was won in the center of the pitch, absolutely dominated by the ball-handling skills of Enzo Fernandez, De Paul and Messi.
- This project demonstrates that some of the critical analysis that some journalists made, were overturned. The final was played at a high rate, and Argentina dominated ball possession by not playing on the counterattack.
- In future works, this network-based approach could be applied to analyze multiple teams or multiple matches, also considering other game situations such as shots on target, game status and opponent quality, offering valuable insights into sports research.





**Thank you for the  
attention!**