Dynamic Analysis of Team Strategy in Professional Football

Dr. Laurie Shaw & Dr. Mark Glickman / Harvard University / laurie.shaw@cfa.harvard.edu

One of the most important game day decisions that a football manager must make is to decide team formation. The choice of formation determines player roles, how they interact, and influences the playing style of both teams during a match. Despite their central role in team strategy, descriptions of formations are still reliant on classifications based on the number of defenders, midfielders and forwards: crude summaries of player configurations that are significantly more fluid and dependent on the game state than '4-4-2' or '3-5-2' would suggest.

We present a new, data-driven technique for measuring and classifying team formations as a function of game state, analysing the offensive and defensive configurations of each team separately, and dynamically detecting major tactical changes during the course of a match. This work used tracking data from 180 matches during a single season of an elite professional league.

Measuring Formations

- Team formations were measured by averaging the position vectors between neighboring players in aggregated local possession windows.
- Defensive and offensive formations were observed separately by combining consecutive periods of possession of the ball for each team into two-minute, non-overlapping windows.
- Figure 1 plots the full set of formation observations for one team during a match. The consistency of the observations indicates that the manager did not make a significant formation change during the match.
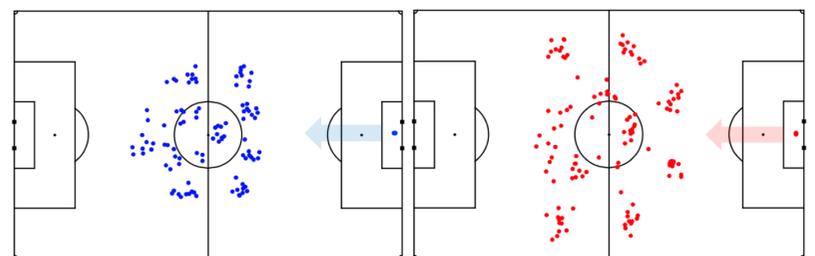

Figure 1 *Left plot*: defensive formation observations, *right plot*: offensive formation observations. Team is orientated to shoot from right to left, and formations are translated to align their center of mass with the center of the pitch.

Identifying unique formations

- Offensive and defensive formations were measured in a training sample of 100 matches, obtaining nearly 8000 formation observations.
- We identified the unique set of formations used in these matches through agglomerative hierarchical clustering, quantifying the similarity of any given pair of formation observations by solving the optimal transport problem using the Wasserstein metric.
- 20 unique formation templates, or clusters, were detected (Figure 2).
- There is a clear ordering to the clusters that highlights the difference between defensive and offensive formations – a distinction lost in previous analyses of formations in football. The top three rows contain predominantly defensive formation observations, while the bottom two rows contain exclusively offensive formations.
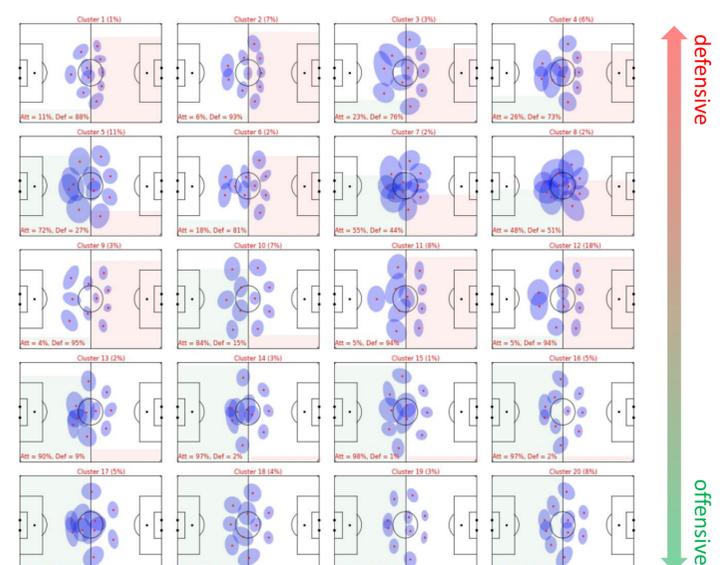

Figure 2 The 20 unique formations identified using hierarchical clustering.

Transitions between defence and offence

- The diagram below gives examples of the defensive and offensive formations that were frequently paired by teams in our dataset.
- In both examples, it is clear how each player would transition between their defensive and offensive roles. The red example indicates that some defensive configurations provide more flexibility in terms of attacking options.
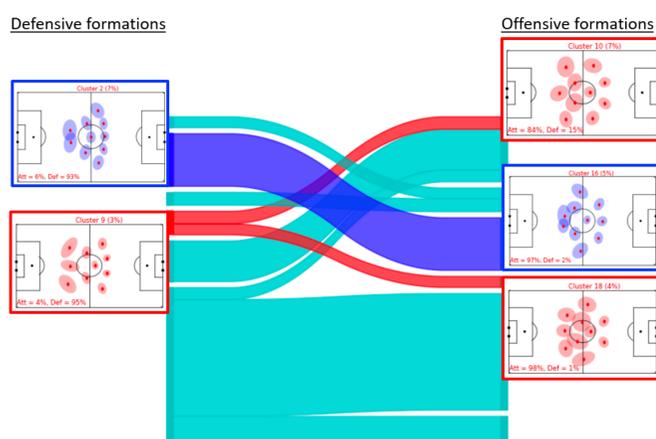

Figure 3 The blue example indicates that teams playing with a defensive formation drawn from cluster 2 (see Figure 2) transition to an offensive formation drawn from cluster 16. The red example indicates that teams that play with defensive cluster 9 transition to either offensive clusters 10 or 18.

Analyzing tactical changes during matches

- Automated detection of formation changes, combined with event data, enable us to investigate why tactical changes were made and evaluate their impact on the outcome of the match (Figure 4).
- The change in defensive formation by the Red team at half time was effective in preventing the Blue team from continuing to create high quality chances from their right side of the pitch, at the expense of conceding opportunities through the centre.
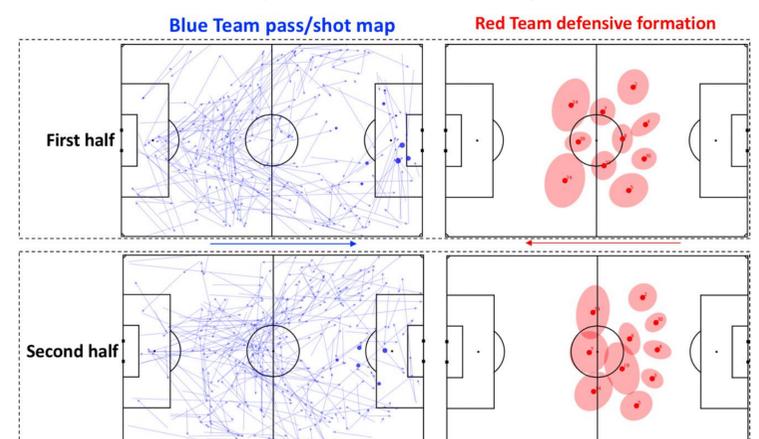

Figure 4 *Right plots:* the change in defensive formation of the Red team before and after half time. *Left plots:* passes (arrows) and shots (circles) of the Blue team (playing from left to right) in the first and second half of the match. Circle size indicates the quality of the shooting opportunity.