Effect of Passing, Possession and Formation in Soccer

Debangan Dey
Johns Hopkins University School of Public Health

and

Soudeep Deb
University of Chicago

Introduction:

In association football, managers decide a team formation to impose his playing style and beat the opposition. Over the years, the game has seen a standard set of formations depending on the position of the players. Moreover, an effective formation depends hugely on the passing and possession abilities of the players.

The motive of this paper is to identify the most effective formation and the best set of players suited for that formation, through analyzing the passing and possession abilities of the players.

Methods:

Throughout this study, we use the data from all the matches in World Cup 2018 provided by StatsBomb (https://github.com/statsbomb/open-data). First, we extract information from 2-minute window just before each shot (we chose 2 minute window as exploratory analyses showed no continuous possession lasting more than 116 seconds). Next, we track where the team won the ball possession that led to the shot, the distance from that location to the shot location, the duration of the movement, the speed of the attack.

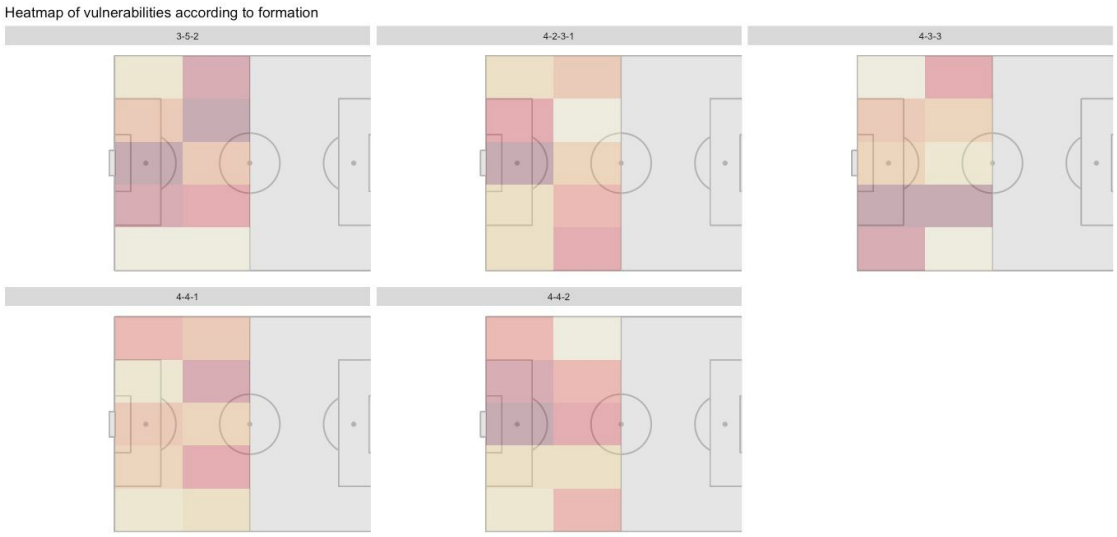
Further, we divide the field into 20 grids of size (30*16) and get the number of successful/unsuccessful passes (scoring/defending team) in each grid during the moments building to that shot.

We perform a detailed logistic regression analysis using goal as a dependent variable and use our extracted covariates. The interesting thing to look at from this analysis will be the effect of the defending team formation, when interacted against the ball win location, on the outcome (goal). Investigating the value of the coefficients across the grid on the field will give us an idea on a particular area of vulnerability of a formation when fielded against another formation.

Once we identify the weak zone for a particular formation, we delve into further investigation around specific players for that position and find association between player level attributes and how the results turned out.

Results:

Using the results from analysis, a preliminary heatmap (*red means more vulnerable*) is generated that helps us identify the most important locations for a particular formation. A quick look suggests 4-3-3 formation is vulnerable through the wings and 4-4-2 is vulnerable through the middle as the team is prone to lose midfield battle. As a next step, we intend to investigate more into the players involved for that position and develop a strategy to field best formation with best players according to an opponent.



Conclusion: This study helps us in identifying the most crucial position corresponding to different formations. Further, this framework equips us to build team strategy for a specific formation and in turn, try to find best players for that particular strategy to do one better than the opposition. Naturally, this method and the results should be very useful for the managers.