Assessing Defensive Players in Football

Student Name : Harleen Gulati , Supervisors: Matthew Penn, Professor Christl Donnelly

University of Oxford

Background and Introduction

Oxford City is a football club which have invested in data of 2 years worth of matches. Plenty of data analysis has been done with this data such as deducing how good of an attacker, passer etc a player is.

Some analysis of defensive ability has been done however more could be done. The current analysis of defensive ability is biased and dependent on the team a player belongs to (if a player belongs to a mostly attacking team, they have less chance to defend and so would be deemed a bad defender which is not necessarily true).

Our objective is to produce defence metrics for players independent of who they play with and against. We intend for our metrics to minimise the following objective function (where AIJ is the ability of player I in season J):

$$\frac{\Sigma_i \Sigma_j \Sigma_k (A_i^j - A_i^k)^2}{\Sigma_i \Sigma_m \Sigma_j \Sigma_k (A_i^j - A_m^k)^2} = \frac{\text{(Variation from players across seasons)}}{\text{(Total variation)}}$$

Note the keywords section at the bottom right explains any unfamiliar football terminology that you may come across.

Methodology

Metric 1

Our first defensive metric:

Split the pitch into 5 20x100 blocks (the pitch is usually 120x80 however the data we use assumes the pitch to be 100x100 - we adjust coordinates in the end to reflect this difference). Assign each block a weight (blocks closer to the goal a player is protecting is scored a higher weight).

The defensive metric of player I in season J is : sum of ((defences done in each block by player I in season J * weight of each block) / total time player I spent defending in season J.)

Note defences here is an umbrella term for interceptions, successful defensive duels and successful aerials.

Metric 2

Defensive metric 2 : for each pass, previous data analysis gives data on the chance of intercepting a pass.

If the pass is intercepted, we are told who made the interception. Split interceptions into two categories : easy to intercept and hard to intercept. Assign each category a weight, with hard to intercept assigned a higher weight.

For successful interceptions, see what the chance of intercepting is. If it is high, add this to the easy to intercept category otherwise add to the hard to intercept category.

Then for player I in season J defence metric 2 is : sum of ((interceptions made in each category by player I in season J * weight of category) / total time player I spent defending in season J).

Metric 3

Defensive duels have four possible outcomes : stopped progress, recovered possession, did both, did neither.

Assign each outcome a weight (e.g. doing both has a higher weight than doing neither) and find how many of each outcome a player did.

Then defensive metric 3 for player I in season J is : sum of ((defences made in each category by player I in season J * weight of each category) / time player I spent defending in season J).

Metric 4

We then compared the objective functions for all metrics but they were mostly similar. Each defence metric focused on a different angle too, so we decided to combine them all.

We assigned each defence metric a weight (based on how much it influences the defending ability of a player) and defence metric 4 for player I in season J is : sum of (weight of each defence metric * score of each defence metric for player I in season J).

Results

Like done with previous data analysis, we combined the scores of both seasons for each player and then we found the percentiles of each players score with respect to other players in the same category (e.g., all the goalkeepers scores were collected and the percentiles found, same with centre backs etc). This makes results easier to understand. A snippet of what is produced is shown below for the player C Haigh

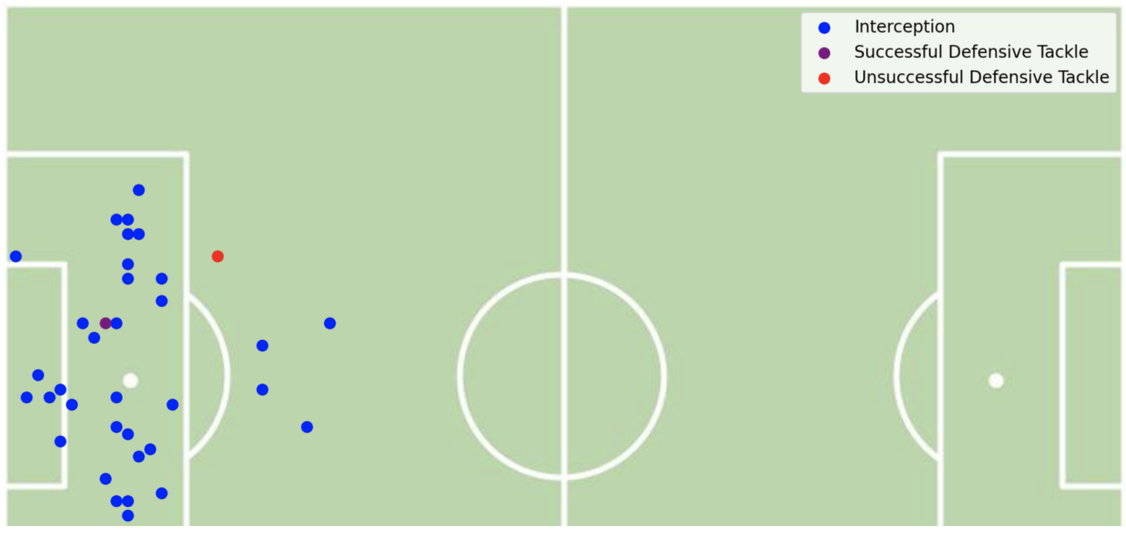
Goalkeepers

	Player Name	Defence Location	Interception Quality	Tackling Quality	Overall
0	C. Haigh	49	13	54	47

Choose team: Oxford City

Choose view: Defending (Visualising)

Choose Player: C. Haigh



Results and discussion

For easier readability the table gives : defence location : 49 , interception quality : 14 , tackling quality : 54 and overall is 47.

The visualisation shown gives us Chris Haigh makes plenty of interceptions close to his goal but only a few tackles (which for a goalkeeper is feasible).

The table does indeed reflect this, for his defence location score is high (most defences closest to his goal). His interception quality seems poor (however the visualisation does not show which interceptions were easy or hard to make) and his defensive duel quality score is 54 which is reasonable since he makes only two tackles, one of which is successful and one which is unsuccessful).

Below we show another table also showing results.

Choose team: Alfreton Town

Choose view: Defending (Score)

Some tables will be empty, this is because there has been no information found for this player from the data used to generate these scores. The data used is from seasons 2021/22 and 2022/23 and so it could be the player did not play in those seasons OR did not play enough hours for data to be collected

Goalkeepers

	Player Name	Defence Location	Interception Quality	Tackling Quality	Overall
0	G. Willis	87	87	69	87
1	J. Askew	62	41	56	58

Centre Backs

	Player Name	Defence Location	Interception Quality	Tackling Quality	Overall
0	S. Brisley	92	49	94	90
1	D. Wiley	96	78	93	90
2	K. Digie	95	59	75	89

Full Backs

	Player Name	Defence Location	Interception Quality	Tackling Quality	Overall
0	H. Perritt	93	96	84	98
1	D. Preston	98	93	89	95
2	J. Clackstone	97	84	86	90

Central Midfielders

Below is the objective function table.

Defence Measure	Objective Function Value
Total defences / time spent defending	0.000183
Defence location / time spent defending	0.000136
Interception quality / time spent defending	0.000148
Defending duel quality / time spent defending	0.000154
Combining all the above / time spent defending	0.000140

Further Research / Next Steps

Tweaking Weights

We tweaked the weights for the defensive duel quality metric and for the interception quality metric using SciPy's minimisation function.

However this minimisation function had issues with the defence location metric and combination metric. A next step could be writing your own minimisation functions to tweak the weights for these metrics so they are optimal in minimising the objective function.

Improving Current Approaches

Our defence location approach splits the pitch into 5 segments. A next step could be looking at splitting the pitch into further smaller segments.

This approach is slightly biased since a goalkeeper will mostly be in the block closest to his goal and so be scored higher than say an attacker.

A next step could be to tackle this bias, by perhaps considering what position a player plays (e.g. defender) and adjusting the score to reflect biases the position may introduce.

Further Metrics

As we have for successful and unsuccessful defensive duels, perhaps we can add a metric for unsuccessful interceptions.

For example, for a pass made which was not intercepted we could see which player of the defending team was closest to this pass. We can then score them down for not taking the opportunity to intercept (here considering to score them down greater if the chance of interception was high).

We can add further metrics like these to provide more information on players defending abilities.

Conclusion

Our metrics seem feasible and reflect the players defence ability fairly. Dividing by the time spent defending reduces biases on the teams ability as a whole in defending or attacking.

However, adding more metrics and tweaking the current metrics could aid in further player defence metrics.

Acknowledgments

Thank you to Matthew Penn for providing all the support, data and advice throughout this project.

Thank you to Professor Christl Donnelly for her support, kindness and advice throughout this project.

Keywords

Interception - a defending player comes in the way of a pass being made between players of the attacking team.

Defensive duels - a player trying to get the ball off an attacker (usually one-one tackling). This either allows the player to stop the progress of the ball (ball goes out of the pitch say) or recover the possession of the ball (the ball now belongs to the defending player) or both or neither.

Aerial duels - ball is in the air and whichever player first touches the ball wins this duel