Using tracking data
Dimitrios Bampalikis
October 2020

Introduction

This report contains the analysis made based on tracking data collected during 2019 from Signality. The dataset provided includes three matches of Hammarby for Allsvenskan, the top Swedish football league. The data consists of x and y coordinates for the players and the ball, extracted from a video in 25Hz frequency (1 observation/frame every 0.04 seconds).

Given this data, we are able to calculate the speed and direction for each of the players, based on the distance covered between the frames. Four different examples of this data can be found in Figure 1.
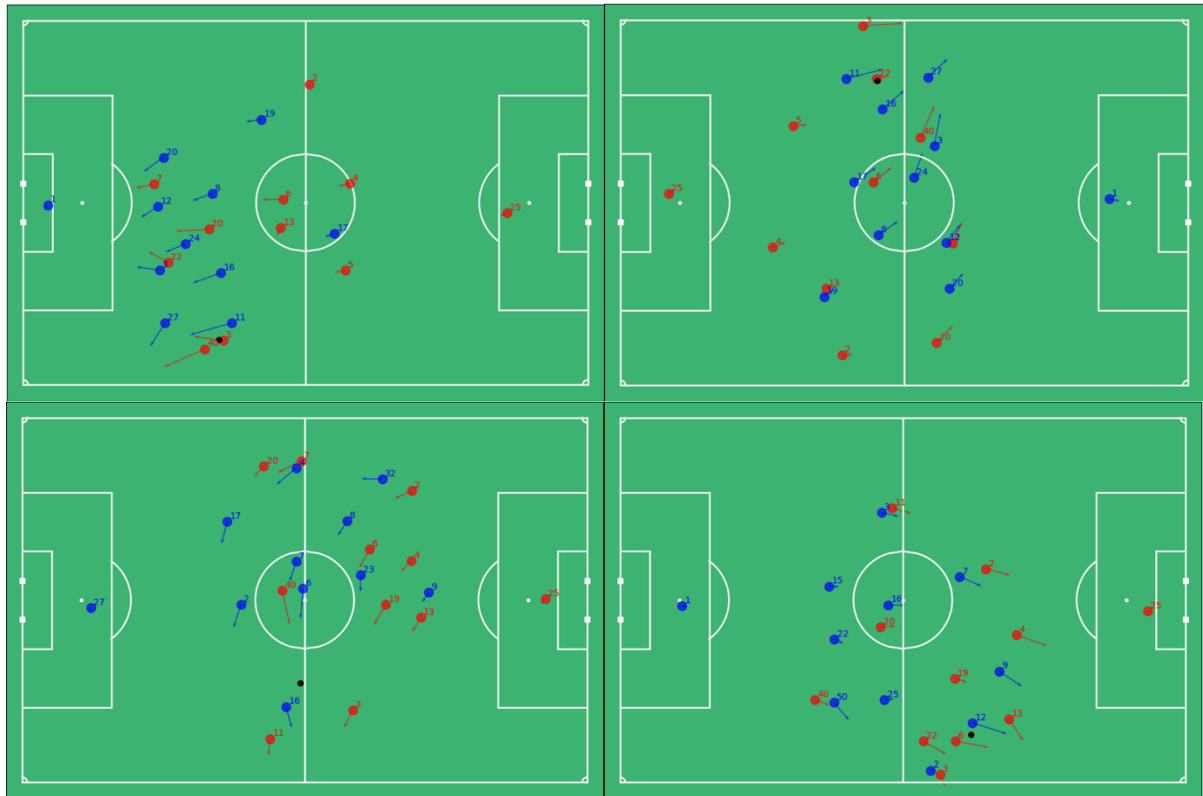

Figure 1: Plots showing the players on the pitch in different timepoints for the three provided matches. The first two (top row) are from Hammarby's game against Elfsborg, while the other two (bottom row) are from the games against Malmö and Örebro respectively. The direction and length of the arrows indicate the direction of the player and the speed at the given moment respectively

Tracking data

Tracking data in principle allows for various different types of analysis. It can be used to extract basic information such as the distance covered, the max speed for a player as well as the distances between the players and their opponents. Furthermore, it enables answering more complex questions, such as the way a team is pressing and where on the pitch does that get initiated, the time it takes to try and recover the ball after losing it. Also, it allows constructing pitch control models for any given moment and evaluating how much different passes contribute to the attacking game, given their probability to succeed. In this last aspect, tracking data combined with event data, facilitates better estimations for the success of an action, given that the position of each player has an impact in that (one example would be a pass and its probability to be successful).

However, there are some limitation when using tracking data, such as the fact that some useful information is not captured and more technical ones regarding the way the data is extracted. Examples of such useful data are the height of the ball at every moment and whether it is curling/spinning (which could have an effect on whether the receiver player will be able to control it), the body orientation of each player (which would help understand how quickly he/she will be able to run towards a specific direction), and therefore the gaze of the players (which could be important in order to evaluate their view of the whole pitch) and more.

The dataset provided has some technical issues such as missing data for the game against Örebro, missing players/ball in some frames, players that "jump" around the field and more. One more limitation of the specific dataset, is the lack of reliable event data, which would make the analysis easier and give further insights to the team's analysis.

In Figure 2, three different variables are plotted for a sequence of 15 seconds for the attacking players of Hammarby: distance from opponents' goal, speed and acceleration. Specifically, given that the tracking data is not totally accurate, the plots include the mean value for each variable in periods of one second. The first two are taken from the 1st and 4th goals against Elfsborg. It is quite interesting to compare these two, since they give a similar image for the three players during these 15 seconds. The last plot shows one of the cases where data is missing.
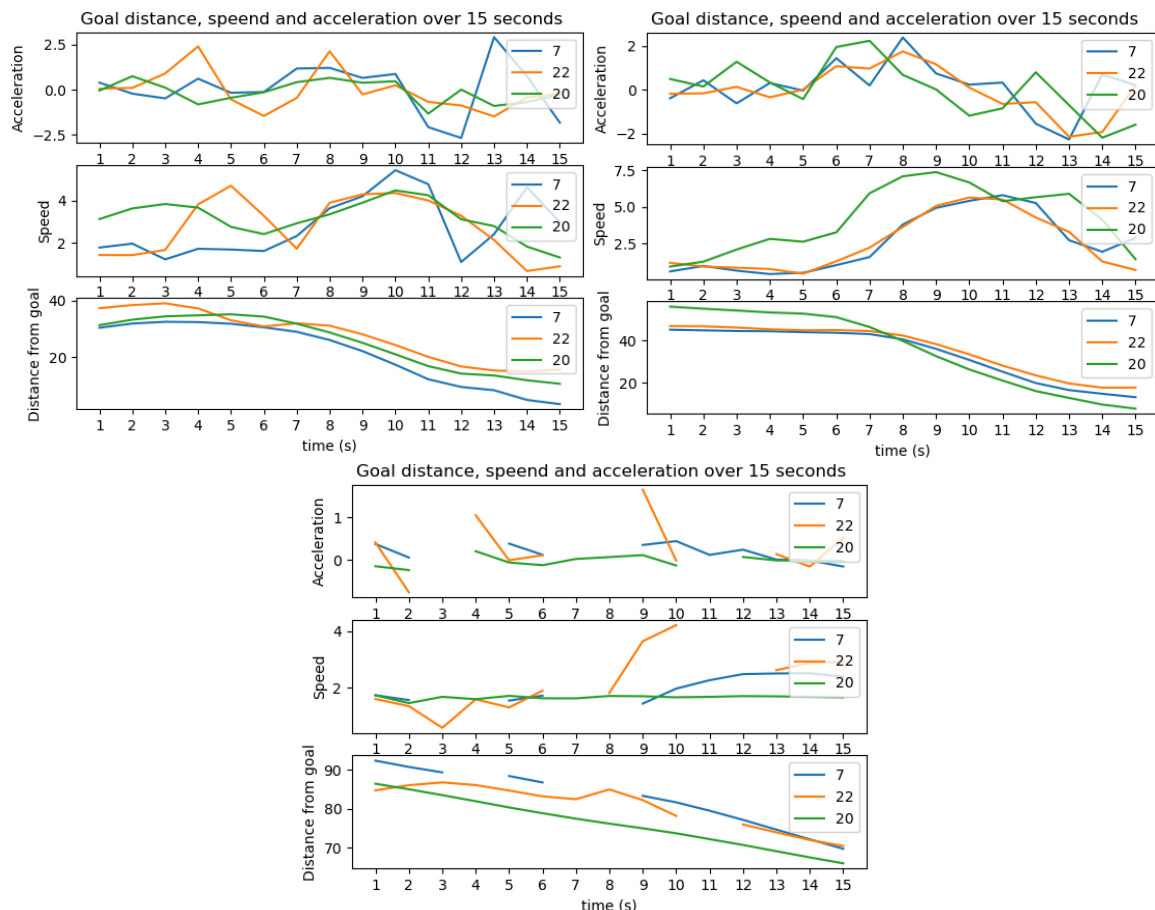

Figure 2: Plots presenting the acceleration, speed and distance from goal for three different players.

Analysing Hammarby's goals

Tracking data enables for analysis related to the players' positions, relative to teammates and opponents. The distance gives a nice overview of the way the team is acting both in the defensive and offensive end. Relatively small distances from teammates would mean more passing options, while bigger distances from the opponents would increase the chance of a pass succeeding. In the same manner, close distance to opponents is the defensive end, would increase the chance of interrupting an attack. The Figure 3a shows the distance of teammates and opponents for five attacking players of Hammarby, in a possession that lead to a goal.

While watching the videos from the games, it becomes quite evident that the left side of Hammarby is responsible for creating many of the opportunities and goals scored. Nikola Durdić (jersey number 40) is making movements with and without the ball, which leads to situations where he is either open for a pass or he forces defenders to follow him, opening the space for his teammates, such as Alexander Kačaniklć (jersey number 20) to make a move in the open space and receive the ball. The latter is achieved also by moving in the opposite direction, but close to one of his teammates, creating confusion to the defender who has to make a quick decision on which player to follow. Figure 3b shows the distance between Durdić and his teammates/opponents for three out of the five goals Hammarby scored against Elfsborg.
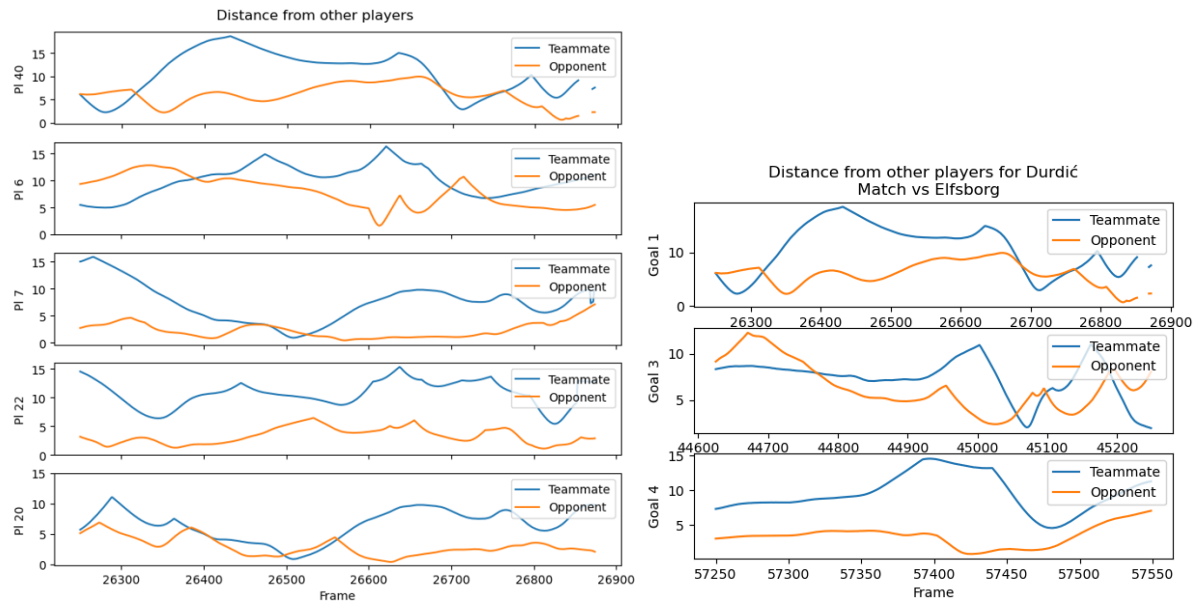
Figure 3: (a) Plot showing the distance between the player in question teammates/opponents. (b) Plots showing Durdić distance from teammates and opponents during three different goals again Elfsborg

Pitch control

Another type of analysis possible with tracking data, is related to the pitch control. Pitch control is the probability that a player would be able to get to the ball when it arrives to a specific location. In other words, the probability that a pass in a specific area will be controlled by a player. To create a model like that, some assumptions are made such as the response time of each player, the speed when running towards the ball. Figure 4 shows plots for the pitch control model for two different matches.

The first plot shows how Hammarby is overloading the left side of the pitch, with Widgren and Durdić, making it impossible for the defender to react. Durdić makes the move to the penalty area, between the defenders, creating the chance for a pass with good probabilities. A similar situation can be observed in the second plot, taken from the match against Malmö. Durdić initially moving in parallel to the goal, makes a move towards it, creating the space for a good pass in the penalty area. Finally, the third plot shows the 4th goal against Elfsborg. In this case too, there are two players on the left side of Hammarby's attack (Durdić and Tanković). Durdić makes the run between the defenders and creates the opportunity for Tanković to move towards the opposite direction, creating enough space for him to shoot and score.

The frames have been selected to emphasize the way Durdić is creating space and opportunities both for passes with good success probabilities, as well as space for his teammates.
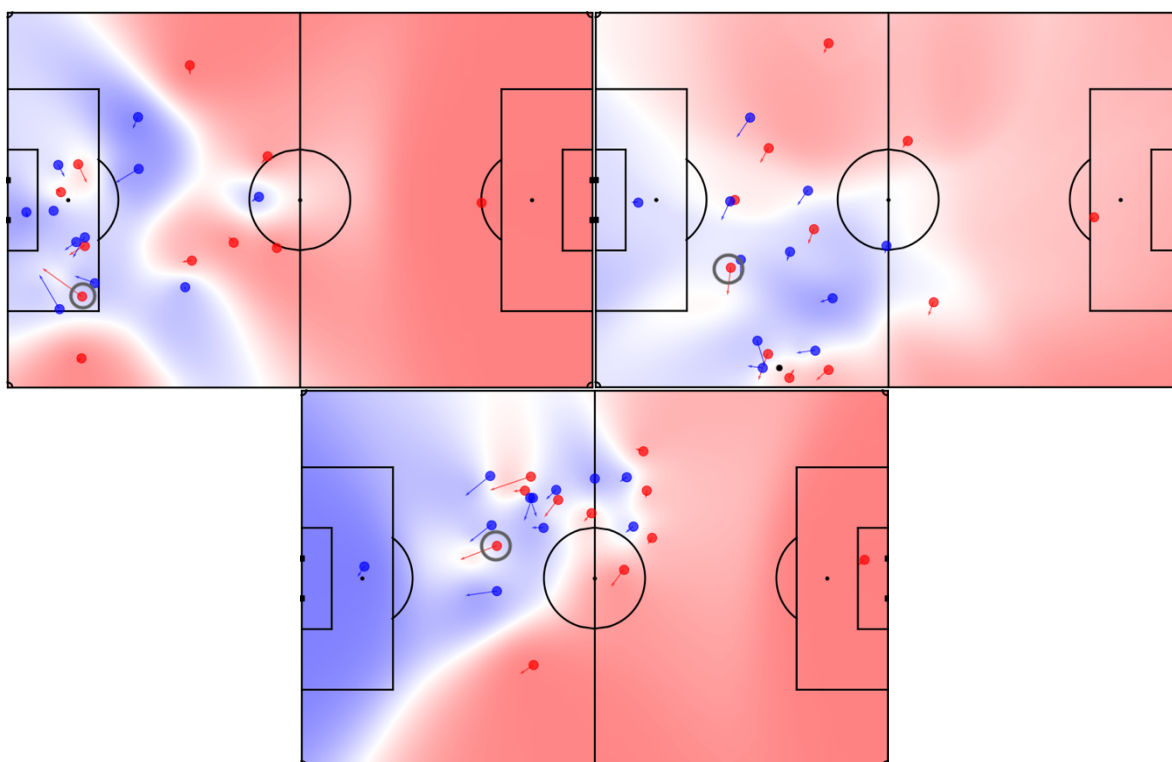
Figure 4: Plots from pitch control model showing the way Hammarby attacks. The left-hand side seems to be one of the driving forces (red areas on that side), while Durdić's (shown with a grey circle) movements open space for passes and teammates.

Notes
- The original sizes of the plots are included in the zip file, since their size had to be reduced in order to make the report smaller.
- The plots were created after editing and adjusting Laurie Shaw's code (and I thank him for it).