Searching for a Strategy: Modelling Player Trajectories in Soccer Games

Jagjeet Singh
jagjeets@andrew.cmu.edu

Rajat Kulshreshtha
rkulshre@cs.cmu.edu

Carnegie Mellon University

Abstract

Traditional methods of predicting outcomes of sports like soccer rely upon averaged statistics like expected goal value and player ratings. While these methods are sufficient to describe the general trend, they are unable to capture the dynamics of human interaction that occur as part of the sport. In this paper, we develop and explore approaches to create representations of the gameplay which are able to predict the trajectory of players. Further, we exploit the spatial relationships between players to model their interactions and improve our predictions.

1. Introduction

Human motion analysis has enjoyed a lot of attention by computer scientists recently, due to a wide variety of important applications which depend upon such analysis. For instance, the task of modeling motion of body parts allows greater insight into creating motion control for robots, and modeling motion of humans relative to each other enables robots that are aware of their surroundings and respond as a human would. Modeling the complexity of human interaction is a challenging task, since the ability of humans to enact complex motion strategies depends heavily upon internalized responses to the environment and surroundings.

Analyzing human motion is a well-established field of active study in robotics and computer vision. It can be broken down into the task of tracking, and subsequently recognizing and predicting the motion given tracking. Of particular interest to our problem is pedestrian detection and tracking. Traditional methods of modeling pedestrians focus on detecting pedestrians [1] as opposed to modeling their motion. Attempts have been made to detect associations within a set of pedestrians [2], which is the precursor to our problem – studying how human interactions affect motion.

The application of such techniques to sports prediction has only recently found traction with the community. Much work has focused on automatic tracking of players [3,4]. Bialkowski et al. continued Peter Carr's work [4] in [5] where they automatically build the heatmap of player positions over the entire match and analyze the effect of team formations in home/away games to determine match outcomes. They divide the playing field into grids which allows for an easier, discrete representation of the game state. Additionally, some groups have been working on modeling the social dynamics between pedestrians [7], but all these approaches are constrained by the need to either handcraft models for interactions, or to label the large datasets with complex annotations marking interactions, which will be prohibitively expensive.

The most relevant work to our project was published by Alahi et al. [6] where they propose a novel data-driven approach to predict pedestrian trajectories. We discuss this further in Section 2. Even so, the study of such behaviors in the context of sports has been very limited. To the best of our knowledge, ours is the first attempt to model the dynamics between players in this specific context.

1.1. Formulation and Motivation

Our aim is to learn a model that is able to i. Learn a good representation of the game state, ii. Model the interactions between the players, specifically between nearby players on the field, iii. Given a sequence of steps, predict the trajectory that the players will follow afterwards. We formulate this as a time series prediction problem over the player coordinates.

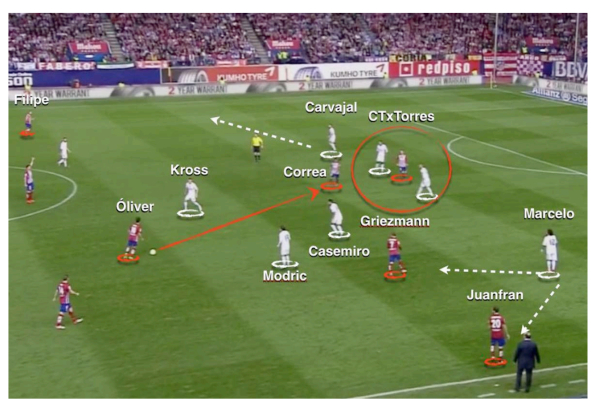

Figure 1: Sample field illustrating playing strategy

Figure 1 shows a sample scenario from a real match. Atletico Madrid (red) is attacking and Real Madrid (white) is defending in this play. The two central defenders have the responsibility of closing the opposing attacker Torres (circled). Also, Correa just moved into an open space to receive the pass from Oliver. Marcelo on left flank, meanwhile, is following a trajectory which makes sure Juanfran and Griezmann don't have an opportunity to run for a through ball. Such decisions are intuitive for a player but very hard for a programmer to model. They are observed very frequently in the game and are an important aspect of good strategies

We aim to learn these behaviors from data. Achieving this goal will allow us to gain insights into the rules that players inherently model through experience, and thus allow us to predict their motion and form strategies accordingly.

1.2. Dataset

The STATS SportsVU dataset [8] is a collection of variable length sequences containing the spatial coordinates of each team's players and the soccer ball. It is equivalent to around 45 matches worth of playing time, and the positions have been sampled at 10Hz. It has been made available to certain research groups on request. We use this dataset in all following discussions.

2. Description of our models

We model the problem as a time series prediction. We describe the representations and models below:

2.1. Game state representations

We use the following game state representations across our models:

2.1.1 Continuous sequences across the field: This is the original format of the dataset where each player and the ball is represented by a set of ordered coordinates over the playing field.

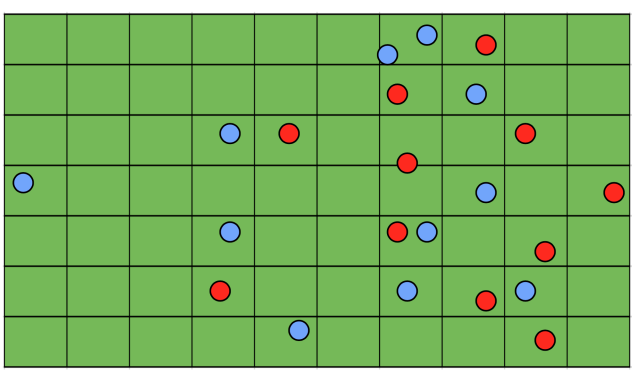

Figure 2: Discrete grid representation.

2.1.2 Discrete grid representation: Since the above representation proved to be difficult to model, we decided to discretize the game state by dividing the field into a grid and indexing each player by the grid they are on. In order to get a reasonable dimension vector while maintaining higher resolution over the field, we decided to pool all the players in each team into a single team vector, and represent the game state by concatenating these two. We decided to ignore the ball in this representation.

2.2. Vanilla Recurrent Model

We first attempted to model the above representations of the game using an LSTM model. After much hyper parametrization, we were unable to train a model that was able to converge or generate reasonable predictions from either representation. We present one such failed outcome below:

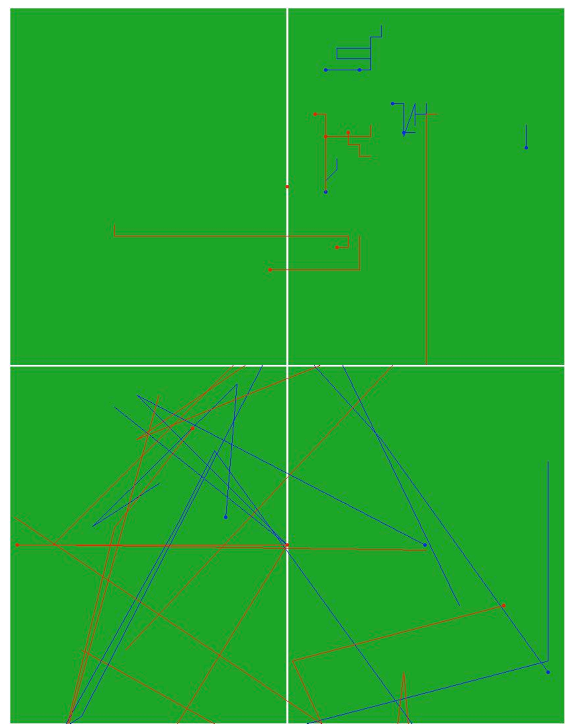

Figure 3: Above: True sequence
Below: Unreasonable outputs from LSTM model

We hypothesize that this might partly be because of the deficiencies in the representations. The discrete representation washes out all of the positional information, since now each grid is just an index. This lack of localization might make it harder for the model to learn the behavior. Further, this makes it impossible for the model to capture the interactions between nearby players, which was a central idea for our task.

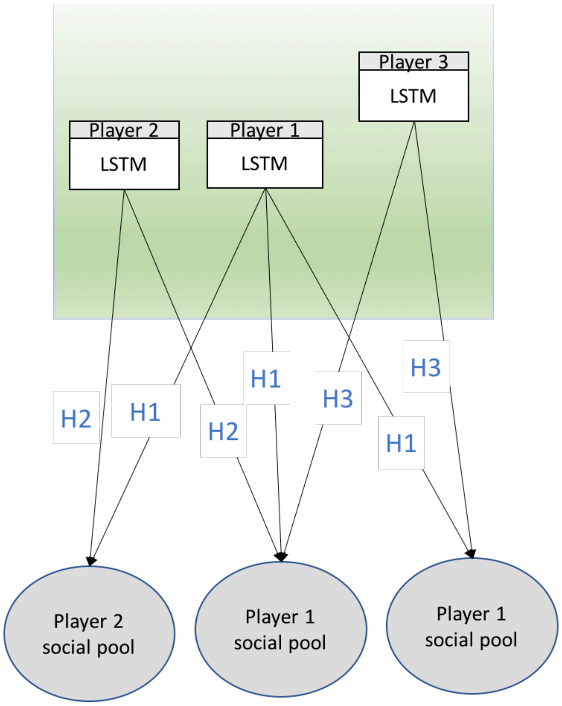

Figure 4: Example of social pooling layer for player 1. Player 2 and Player 3 are the only players in the neighborhood of Player 1. As such, social pooling layer of player will have inputs of hidden states of players 2,3 and 1.

2.3 Social-LSTM

In Section 2.2, we explained how vanilla-LSTM tried to predict trajectories of individual players on the field. The caveat associated with it is that it fails to capture the spatial interactions of different players in the same temporal frame. This means that it will try to predict the next position of each player without considering the effect pf other players in the neighborhood.

This problem gained recognition recently and there has been growing interest in analyzing how humans respond to other humans around them. Most of the work in this domain has focused on pedestrians. [6] proposed the Social-LSTM model to capture the behavior of humans waking in densely crowded environment. Humans have the tendency to change their path if there are other humans on the way, prefer open spaces rather than small openings even if that leads to a longer distance and other such behavioral attributes. Such actions are more of an intuition for humans and are difficult to model. This paper tried to adopt a similar approach for soccer players, instead of pedestrians.

During a soccer match, there are a lot of strategies that players adopt. Defenders try to close in the opposing attacker to increase the possibility of intercepting the pass. Attackers try to move into open spaces to receive the pass. Players try to predict the trajectories of other players and decide their movement accordingly. Similar to the case of pedestrians, such behavioral attributes are difficult to model and can mainly be learnt from the data.

2.3.1 Implementation

This section also uses the data provided by STATS as described in section 1.2. At any time frame, each of the 22 players and the ball are represented by their xy-coordinates. We observe the positions for the first few time frames and predict the positions for next few time frames. The model consists of LSTMCell for each player and the ball.

The hidden state of each LSTM has the capacity to capture time-varying motion properties of the corresponding player. Figure 4 shows how social pool layer for player 1 is impacted by other players in the neighborhood. This approach is however insufficient to keep track the spatial distance of players in the neighborhood. For example, in Figure 4, the spatial distance between player 3 and player 1 is more than that between player 2 and player 1. The model however won't realize that and will give equal weight to their hidden states.

To overcome this problem, occupancy matrix is created. The neighborhood for each player is divided into a grid. The hidden states for players in the same grid are stored at the same index in the occupancy matrix as shown in Figure 5. The hidden states are then fed to the social pooling layer using a weighted sum as described below:

$$H_t^i(m,n,:) = \sum_{j \in \mathcal{N}_i} \mathbf{1}_{mn}[x_t^j - x_t^i, y_t^j - y_t^i] h_{t-1}^j, \qquad (1)$$

where index j is for the player in the neighborhood of player i and $\mathbf{1}_{mn}$ [x,y] is an indicator function to check if (x,y) is in the (m,n) cell [6].

For predicted position, it is assumed that coordinates follow a bivariate Gaussian distribution. The hidden states at time t are used to get 5 parameters which include mean of x and y, standard deviation of x and y and correlation coefficient. The coordinates are then sampled from the resulting Gaussian distribution.

Loss function in Social LSTM is the negative log likelihood of x and y coordinates, given mean, standard deviation and correlation coefficient.

2.3.2 Results

Figure 6 shows the results on a sample sequence. The trajectories of the players are observed for a few time steps (solid lines). Predicted trajectories are then plotted (dashed lines) as a continuation on the same chart. As can be seen in the figure, Defenders D1 and D2 are initially closing down the attacker A1 in the observed part of the sequence.

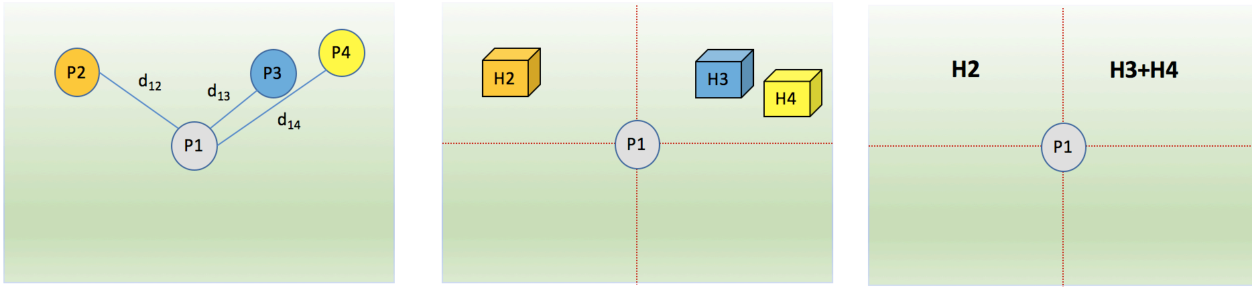

Figure 5: Occupancy matrix for player 1's neighborhood. As player 3 and 4 lie in the same grid, their hidden states will be summed over and saved in the corresponding index of the occupancy matrix of player 1.

But, as soon as the attacker A1 sees an open space, it changes the path and tries to occupy that space. This however alerts the defender D2 and as such he also follows the attacker A1 to close him down. Meanwhile, defender D1 who was earlier covering A1 now finds another attacker A2 in its neighborhood. As such, D2 changes path to close that defender down.

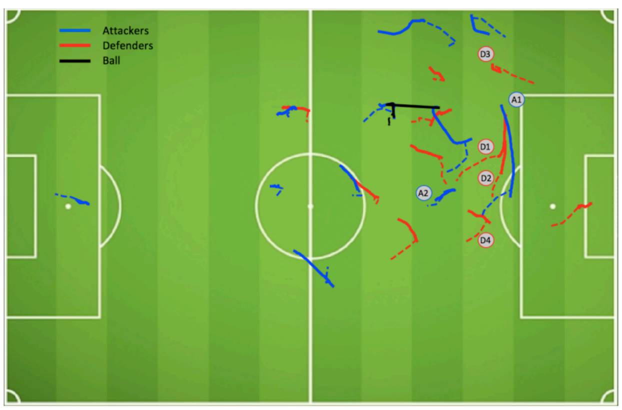

Figure 6: Results on a sample trajectory from a real game

2.3.3 Fallbacks

The model worked well for smaller sequence lengths (of the range 20). However, as the length increased, it diverged from the true positions significantly. It started failing when the sequence length reached 100.

Another shortcoming of the model is that it treats each player equivalently. However, in a real game, a defender might react differently depending on whether the player in the neighborhood is a fellow defender or the opposing attacker. That said, we believe that such scenarios are implicitly considered in our case because of the data format. The index for defending team (1-11) and attacking team (12-22) don't change from sequence to sequence. Also, unlike the case of pedestrians, the number of players don't change over time.

3. Conclusion

We presented an approach for LSTM-based trajectory prediction in a soccer match. Although trajectory prediction is a common research focus but applying it sports domain is not often seen. We started with the importance of analyzing trajectories of players in a real match. The problem of trajectory prediction was taken as a case of sequence prediction. We first tried a vanilla-LSTM based approach and found the results unsatisfactory. Then Social-LSTM was used to consider interaction of players who are in different spatial position in the same time frame. We discussed the implementation and how this case differs from the original use case of pedestrian trajectory prediction for Social-LSTM. Finally, applied the model on some sample sequences wherein we observed some portion of the sequence and tried to predict the remaining sequence. The results were found to be intuitive in many such predictions.

References

[1] Dollar, Piotr, et al. "Pedestrian detection: An evaluation of the state of the art." IEEE transactions on pattern analysis and machine intelligence 34.4 (2012): 743-761.
[2] Pellegrini, Stefano, Andreas Ess, and Luc Van Gool. "Improving data association by joint modeling of pedestrian trajectories and groupings." European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2010.
[3] Needham, Chris J., and Roger D. Boyle. "Tracking multiple sports players through occlusion, congestion and scale." BMVC. Vol. 1. No. 1, 2001. 2001.
[4] Liu, Jingchen, et al. "Tracking sports players with context-conditioned motion models." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2013.
[5] Bialkowski, Alina, et al. "Win at home and draw away: Automatic formation analysis highlighting the differences in home and away team behaviors." Proceedings of 8th Annual MIT Sloan Sports Analytics Conference. 2014.
[6] Alahi, Alexandre, et al. "Social lstm: Human trajectory prediction in crowded spaces." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
[7] Koppula, Hema S., and Ashutosh Saxena. "Anticipating human activities using object affordances for reactive robotic response." IEEE transactions on pattern analysis and machine intelligence 38.1 (2016): 14-29.
[8] Data Source: STATS SportsVU. www.sportvu.com., copyright 2017.