Analyzing and Predicting Events in Soccer and Tennis using Spatiotemporal Data

X. Wei, S. Sridharan
Queensland. Uni of Tech
Brisbane, Australia

P. Lucey, R. Navarathna
Disney Research Pittsburgh
Pittsburgh, PA, USA

S. Morgan
Australian Institute of Sport
Canberra, Australia

ABSTRACT

In this paper, we summarize our recent work in analyzing and predicting behaviors in sports using spatiotemporal data. We specifically focus on two recent works: 1) Predicting the location of shot in tennis using Hawk-Eye tennis data, and 2) Clustering spatiotemporal plays in soccer to discover the methods in which they get a shot on goal from a professional league.

1. INTRODUCTION

Over the last 10 years, there has been an enormous interest in using data to help drive decisions in sports. However, nearly all of the analytical works have dealt solely with hand-labeled event data which describes *what* happened (e.g. basketball - rebounds, points scored, assists, football - yards per carry, tackles, sacks, soccer - passes, shots, tackles). As most sporting environments tend to be dynamic with multiple players continuously moving and competing against each other, simple event statistics do not capture the complex aspects of the game. To gain an advantage over the rest of the field, sporting organizations have recently looked to employ commercial tracking technologies which can locate the position of the ball and players at each time instant in professional leagues (e.g., Prozone, STATS SportsVU, Hawkeye) - to determine *where* and *how* events happen. In this paper, we summarize our recent work in using spatiotemporal tracking data to analyze and predict events in sports. Specifically, we highlight our recent work in predicting the next shot location in tennis, as well as discovering methods of how teams score in soccer by first aligning and then clustering player's trajectories.

2. PREDICTING SHOTS IN TENNIS

Our goal of this work was to be able to predict the location of shot that would be played in tennis. As a player tends to have their own individual "signature", the motivation was to model each player's individual behavior via data. To do this we used an entire tournament of spatiotemporal data from

WOODSTOCK '97 El Paso, Texas USA

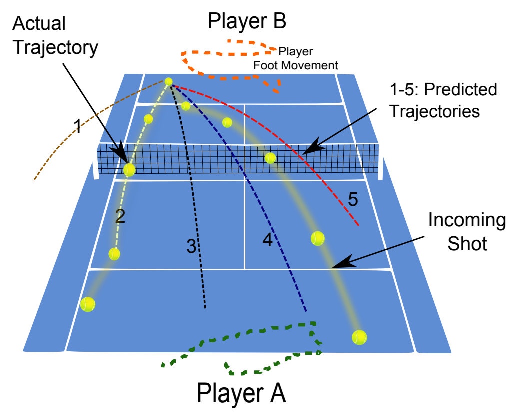

Figure 1: The tennis tracking data contains both ball trajectory and player feet movement information. In this example, Player A (bottom) serves the ball to Player B (top) who returns the ball to Player A. We are interested in predicting the location of this next shot based on the incoming shot (Dashed lines indicate possible trajectories of next shot).

Hawk-Eye [1] from the 2012 Australian Open (see Figure 1). This task is much more challenging than the previous work of Wei et.al. [7], where they predicted "what" type of shot (i.e., winner, error or continuation) but not "where". Having the ability to predict the location of future short-term events has enormous potential in area of automatic broadcasting, coaching and analysis.

For this work, we used Hawk-Eye data which records the (x, y, z) positions of the ball as a function of t. For each point in a match, we also have additional meta-data such as shot type, outcome, server, score, etc. An example of the ball trajectories in a point is given in Figure 1. Player court positions are recorded as the (x, y) positions of players on the court at 20 frames per second. By leveraging these tracking data, we are able to augment any point of interest with detailed ball trajectories and player court positions over time. For this work, we used the data from the 2012 Australian Open Men's draw which consisted of more than 10,000 points. We specifically modeled the behavior associated with the top 3 seeds at the tournament (Novak Djokovic, Rafael Nadal and Roger Federer) as they had the most data and it also allowed us compare the different styles of play (Nadal is left handed while Federer and Djokovic are right handed.)

Table 1: Description of the shot variables used in this paper.

Feature	Description
Speed	Shot average speed
Angle	Angle between shot & center line
Feet Location	Player and opponent court position when shot starts
Shot-Start Loc.	Location where shot starts
Shot-End Loc.	Location where shot impacts the court
No. of shots	Total number of shots in the point
Opponent Movement	Local speed & direction of the opponent before the player strikes the ball

2.1 Features

Before generalizing a player's behavior, we need to find features/factors to represent each shot. Given a shot's start and end time, we are able to find its starting and ending location. Using this information, we calculate the angle, maximum height, average speed and instantaneous speed of the shot. To add player information, we calculate the court position for both players at start and end time of the shot. Table 1 presents a summary of the shot features used in this paper.

2.2 Model

To model and predict shots over time, we tested our method on two models:

2.2.1 *Dynamic Bayesian Network*

A Dynamic Bayesian Network (DBN). The DBN framework has two levels: 1) a Bayesian Network (BN) that captures varying factors of a rally, and 2) the temporal aspect using the previous state information which can be obtained via the 2-timeslice Bayesian Network (2TBN).

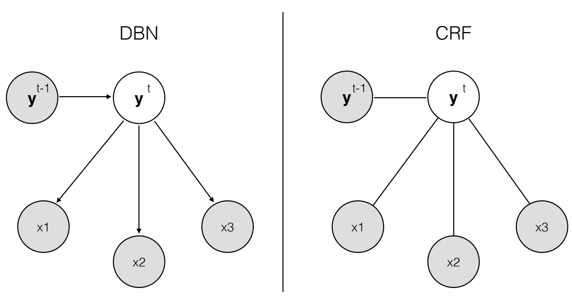

Figure 2: (Left) We use a DBN to predict the next state $\mathbf{y}^t$ given the current observation $\mathbf{x}^t$ and previous state $\mathbf{y}^{t-1}$(gray nodes are observed and clear nodes are hidden). Here y is a discrete label representing the region of a shot location (b) We use a Conditional Random Field to predict the next state $\mathbf{y}^t$ given the current observation $\mathbf{x}^t$ and previous state $\mathbf{y}^{t-1}$.

Given that our observation or feature vector $\mathbf{x}^t$ contains information about the incoming shot (i.e. location, speed, angle, player's feet position, number of shots in rally, etc) at time t, and we know the previous state $\mathbf{y}^{t-1}$ (i.e. previous shot location), using our model topology shown in Figure 2(b), we can infer the probability of next shot landing at each zone by using Baye's law:

$$P(\mathbf{y}^t|\mathbf{y}^{t-1}, \mathbf{x}^t) = \frac{P(\mathbf{x}^t|\mathbf{y}^t)P(\mathbf{y}^t|\mathbf{y}^{t-1})}{P(\mathbf{x}^t|\mathbf{y}^{t-1})} \tag{1}$$

where the next state is conditioned on the previous state. Our prediction is the state, $\mathbf{y}_i^t$, with the highest probability.

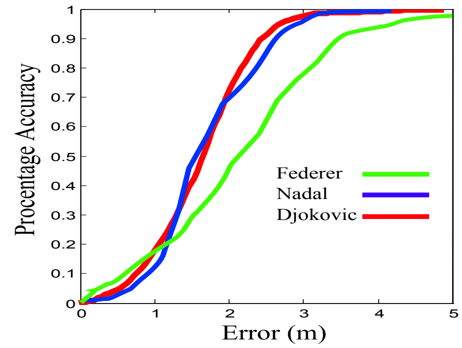

Figure 3: Plot shows the percentage accuracy against error distance in meter to measure the performance of continuous prediction.

Table 2: Prediction percentage accuracy for different quantization levels

Levels	1	2	3	4
DBN	77.3%	44.2%	30.4%	12.7%
CRF	79.2%	61.5%	41.2%	30.1%

2.2.2 *Conditional Random Field*

In the previous model, DBN, features are assumed to be independent. It is a generaltive model. However, this is not the case in practical as many shot factors are related to each other. Therefore, we also model the player behaviour using a Conditional Random Field (CRF). A CRF relaxed the feature independent assumption by directly modelling the conditional probability. Given the observation and previous state, the probability of next shot landing at each zone is:

$$P(y_t|\mathbf{x}) = \frac{1}{Z}\{\Psi_1(y_t; X, \theta_1) + \Psi_2(y_t, y_{t-1}; X, \theta_2)\}, \tag{2}$$

where y_t is the landing location of the shot at time t. θ_1 is a set of parameters that correspond to feature X and states $\mathbf{y}$. θ_2 is a set of parameters that correspond to feature X and edges between y_t and y_{t-1}.

We define Ψ_1 and Ψ_2 as:

$$\Psi_1(y_t; X, \theta_1) = \exp\{\theta 1^T X\} \; and \tag{3}$$

$$\Psi_2(y_t; X, \theta_2) = \exp\{P(y_{t+1}|y_t; X, \theta_2)\}, \tag{4}$$

Z is a normalization constant.

2.3 Results

We tested our model against different quantization levels. When level equals one, we only consider two possible outcomes: inside and outside. When level is two, we quantize the inner court into two regions. Results can be found in Table 2. We also check the predictability for each players (See Figure 3). For full descriptions of these works, see [6, 7]

Figure 4: An example of the tracking data for both player and ball. Player positions are shown with their assigned role (discussed in Section IV.A.). In this work, we focussed on a team which used a 4-3-3 formation.

3. CLUSTERING SOCCER PLAYS

Our goal is to discover the most likely methods a team uses to score in soccer. This is essentially a clustering task, however, due to the constant swapping of players the trajectories need to be aligned. In this section we summarize our work in this area.

3.1 Data

To enable our research, we utilized the (x, y) positions of both players and ball across 9 complete matches (14 hours of tracking data) from a top-tier professional soccer league (See Figure 4 for an example). The fidelity of the data is at the centimeter level, and was sampled at 10 fps. In each of these 9 matches, one team was constant while the opposition was different. These matches also had associated event-level data. For our analysis, we focussed our analysis on the one constant team (i.e., our analysis was independent of the opponent).

3.2 Aligning Trajectories

The most troubling issue is that of ”permutations”, and forming a feature representation which can maintain the feature correspondence. Such representation has to deal with player substitution and the explosion of permutations permutations. Recently, researchers have looked to using a “role-representation” [4] to analyze team behavior. A role representation is more meaningful since formations are defined by roles and individual responsibilities, not identities (e.g. names or jersey numbers). For example, in a 4-3-3 formation, roles are defined as $\mathcal{R} = \{$left back, right back, left center back, right center back, left center midfield, right center midfield, left wing, right wing, attacking center midfield, striker, goal keeper$\}$

Given player positions $\mathbf{p}_t^\tau = [x_1, y_1, x_2, y_2, ..., x_P, y_P]^T$ of P players in team τ at each time instant t. The goal here is to find the permutation matrix, $\mathbf{x}_t^\tau$, which give us a vector of role positions: $\mathbf{r}_t^\tau = \mathbf{x}_t^\tau \mathbf{p}_t^\tau$. Note here, each element $x_t(i, j)$ is a binary variable, and every column and row in x_t must sum to one. If $x_t(i, j) = 1$ then player i is assigned role j.

In [4], a four stage[1] approach is proposed to tackle the problem. First of all, a-state-of-the-art player detector is employed to detect player positions at each timestamp. For each observation, it provides player position (x, y), a timestamp t and a team affiliation estimate $\tau \in (\alpha, \beta)$. If any player is miss detected, an algorithm will be used to infer the position of that player based on spatiotemporal correlations.

The role assignment task is formed as an optimization problem where the goal is to minimizes the square L_2 reconstruction error.

$$\mathbf{x}_t^{\tau *} = \underset{\mathbf{x}_t^\tau}{\operatorname{argmin}} ||\hat{\mathbf{r}}^\tau - \mathbf{x}_t^\tau \mathbf{p}_t^\tau||_2^2 \quad (5)$$

This is a linear assignment problem where the cost of each entry is:

$$\mathbf{C}(i, j) = ||\hat{\mathbf{r}} - \mathbf{p}_t(j)||_2 \quad (6)$$

To solve the assignment problem, the mean formation is used for initialization as the team should maintain this basic formation in most circumstances. In [4], the mean formation is found by taking the mean value of $200,000$ frames of human annotated data. Finally, the optimal solution can be found using the Hungarian algorithm [2]. For more details, please see [8].

3.3 Clustering

Once the representation is resolved, the next challenge is to find meaningful patterns from the large volumes of data which can uncover common patterns of a team's play. For soccer, the clear objective of a team is to score more goals than the opposition. Even though other latent variables are present (i.e. passing patterns), event such as shots and corners are probably the most important events to analyze.

To tackle the problem, we segment goal-scoring opportunities by going back 10 seconds from the shot. After the data has been segmentation, we employ to an unsupervised clustering method to uncover patterns of a team's play. K-Means algorithm is a widely used method for unsupervised clustering. In [4], K-means is used to find the top N formations or tactics in Hockey. However, K-means clustering requires a good initialization. If K is not chosen properly, dissimilar data may be clustered into one cluster. To avoid this situation, we use a hierarchical clustering method which is similar to matching-pursuit but based on examples and not a linear combination of examples[2].

In the beginning, we treat each sample as a cluster. The algorithm can then be described as following:

1. Randomly select a cluster;
2. Compute the distance between each clusters;
3. If there exist at least one cluster with a distance less than the threshold t, go to step 4; else go to step 7;
4. Find the nearest cluster from the selected cluster;
5. Merge two clusters and update the centroid k, this merged cluster become new selected cluster;
6. Repeat step 3-6;

[1]The first two steps are skipped in this research since we are dealing with tracking data

[2]Due to rules of the game, a linear combination may result in situations which can not exist so specific match examples are preferred.

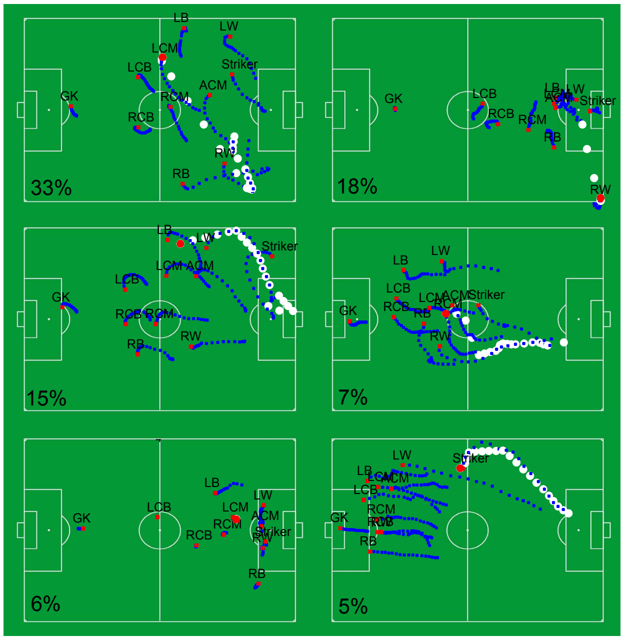

Figure 5: Plot showing top six scoring patterns of a team. Percentage on the left bottom indicates the how often a team score in this way. Red dot indicates the starting location of each player/ball.

7. Store the merged cluster, remove this merged cluster from the set;
8. If there is any cluster left, repeat step 3-7; else finish clustering;

3.3.1 *Offensive Play Analysis*

To find a team's most probable methods of scoring, our clustering method is performed on all goal-scoring opportunities for this team which includes normal shots, corners and free-kicks. Figure 5 shows the top six scoring methods for a team. The red dot indicates the starting location of each player/ball. As can be seen in this figure, in the first example shown in the top-left corner, 33% of the goal scoring opportunities occur in this fashion where the ball starts on the left-hand-side on the half-way line and then moves to the right-wing who cuts back to the top of the box for a shot on goal. The second top method of getting a shot on goal is via a corner-kick from the right hand side (18%). The third top method starts again from the left-back starting from the half-way line. The fourth and sixth methods seems to be counter attacks, while the fifth is via a free-kick.

3.3.2 *Corner Analysis*

Not only can continuous plays be analyzed, but also set-pieces such as corner-kicks. Figure 6 shows clustering result for all corners. In 1-4, the ball is kicked directly to a striker, who is trying to deflect the ball by head or foot into the goal. 5 and 6 both look like set plays, designed to draw out the defense away from the goal mouth, possibly to create more space for an eventual goal shot. As can be seen in most situations, the team of interest has their defenders around the center of the field which means the flank are potential outlets for a quick counter attack.

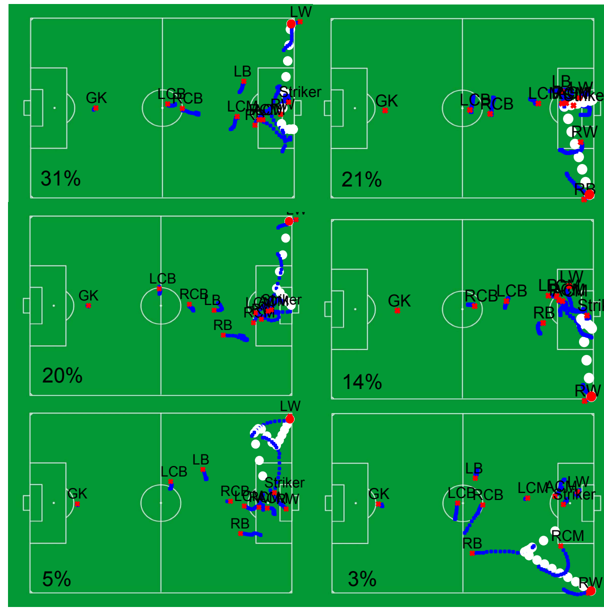

Figure 6: Corner Analysis.

4. CONCLUSIONS

In this paper, we have highlighted our recent work in analyzing and predicting behaviors in sports using spatiotemporal data focussing on predicting the next shot location in tennis as well as clustering plays in soccer. For more details, please refer to the following papers [7, 8, 6, 3, 4, 5].

5. REFERENCES

[1] Hawk-Eye. `http://www.hawkeyeinnovations.co.uk`.

[2] H. W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1-2):83–97, 1955.

[3] P. Lucey, A. Bialkowski, P. Carr, E. Foote, and I. Matthews. Characterizing multi-agent team behavior from partial team tracings: Evidence from the english premier league. In *AAAI*, 2012.

[4] P. Lucey, A. Bialkowski, P. Carr, S. Morgan, S. Morgan, I. Matthews, and Y. Sheikh. Representing and discovering adversarial team behaviors using player roles. In *CVPR*, 2013.

[5] P. Lucey, D. Oliver, P. Carr, J. Roth, and I. Matthews. Assessing team strategy using spatiotemporal data. In *ACM SIGKDD Conference on Knowledge, Discovery and Data Mining (KDD)*, 2013.

[6] X. Wei, P. Lucey, S. Morgan, and S. Sridharan. Predicting shot locations in tennis using spatiotemporal data. In *Digital Imaging Computing: Techniques and Applications*, 2013.

[7] X. Wei, P. Lucey, S. Morgan, and S. Sridharan. Sweet-spot: Using spatiotemporal data to discover and predict shots in tennis. In *MIT Sloan Sports Analytics Conference*, 2013.

[8] X. Wei, L. Sha, P. Lucey, S. Morgan, and S. Sridharan. Large scale analysis of formations in soccer. In *Digital Imaging Computing: Techniques and Applications*, 2013.