Automatic Event Detection in Basketball

Suraj Keshri, Min-hwan Oh, Sheng Zhang, Garud Iyengar
September 22, 2017

IEOR Department, Columbia University
Introduction
Motivation

- Recognizing player **match-ups and game events**, e.g. ball screen, drive, post-up, etc., crucial for gaining insights both on players and teams
- Manually labeling these events not scalable
- **Goal**: Detect events automatically using player tracking data!
Installed in 2013. Tracks:

- $(x, y)$ locations of all 10 players
- $(x, y, z)$ locations of ball
- 25 observations per second
- Event annotations (shots, passes, fouls, etc.)

1230 games per season: $\approx 1$ billion space-time points per season
Defense Assignment
Defense Attraction in Basketball
Basic Setting

- $D_{ti}$: location of defender $i$ at time $t$
- $O_{tj}$: location of offender $j$ at time $t$
- $l_{tij} = 1$: $i$ guards $j$ at time $t$

Stochastic model

$$D_{ti} | l_{tij} = 1 \sim \mathcal{N} \left( Z_{tj}^T \Gamma, \sigma_D^2 \right)$$

[Franks et al. (2015)]

- Defender location is determined by offender characteristics
Basic Setting - player and location dependency

• Γ is player and location dependent:

\[
\Gamma_{pk} = \begin{bmatrix}
\gamma_{pk}^o, \\
\gamma_{pk}^b, \\
\gamma_{pk}^h,
\end{bmatrix} \\
p = g(\cdot, \cdot) \text{ grid picker}
\]

\[
D_{ti|tij} = 1 \sim \mathcal{N} \left( Z_{tj}^T \Gamma g(t, j), \sigma_D^2 \right)
\]

• Prior on \( \Gamma_p = [\Gamma_{p1}, \ldots, \Gamma_{pK}] \)

\[
\Gamma_p \sim \mathcal{N} (\mu_{\Gamma}, \mathcal{K})
\]
• Model the evolution of man-to-man defense using HMM
• Hidden states ($I_t$): defensive mapping

- $I_0$ → $I_1$ → $I_2$ → $I_3$ → …
- $D_0$ → $D_1$ → $D_2$ → $D_3$

• How about transition probability?
Transition Probability

- Total of $5^5 (= 3125)$ matchings $\Rightarrow$ intractable to learn probabilities for all transitions
- Propose a bond energy based defensive assignment transition
  - Single defensive match-up: bond
  - Defensive switching: breaking and forming a new bond.
- 4 types of bonds: 1-on-1 on-ball (or off-ball) bond, extra on-ball (or off-ball) bond
- Transition probability proportional to energy difference

\[ P(I_t \rightarrow I_{t+1}) \propto e^{-\Delta E_{t,t+1}} \]
· Double team match-ups have higher energy (more unstable) than 1-1 match-up. Hence, more likely to switch to 1-1 match-up.

**Figure 1:** On-ball double team

**Figure 2:** One-to-one match-up
• For player $p$ and location $k$: sample 
\[ \Gamma_{pk} \sim \mathcal{N}(\mu_\Gamma, \mathcal{K}) \]
• For all times $t$, sample defensive assignment $l_t$ using energy based transition

\[ D_{ti}|l_{tij} = 1 \sim \mathcal{N}(Z_{tj}^T \Gamma_{g(t,j)}, \sigma_D^2) \]

• Iterate until convergence
• Initialize all the fixed parameters for \( GP \) prior, bond energies \( e \), and \( \sigma_D^2 \). Let \( \theta \) denote all the fixed parameters.

• Until convergence
  • Sample from \( P(I|\Gamma, D, \theta) \) using forward filtering backward sampling algorithm
  • Update energy parameters \( e \) given the sample of \( I \)
  • Sample \( P(\Gamma|I, D, \theta) \)
  • Update kernel parameters, and \( \sigma_D^2 \) given the sample of \( \Gamma \)
Event Detection
Event Detection

- Want to detect events *without* labeled data
- Model sequence of event progression using HMM
- Define the binary hidden state at each time point as an indicator of whether an event is taking place or not
- Specify the parametric form of the emission distributions which are characteristic to actions
- Using HMM, compute most likely sequence of hidden state
Ball Screen

- $S_t$: indicator of ball screen event
- $X_t$: distance between on-ball defender and potential screener
- $Y_t$: distance between hoop and ball handler
- $W_t$: speed of potential screener

$$X_t|S_t = 1 \sim \exp(\lambda_x)$$
$$Y_t|S_t = 1 \sim \log\mathcal{N}(\mu_y, \sigma_y^2)$$
$$W_t|S_t = 1 \sim \exp(\lambda_w)$$
$$X_t|S_t = 0 \sim \log\mathcal{N}(\mu_x, \sigma_x^2)$$
$$Y_t|S_t = 0 \sim \text{Unif}(0, \theta_y)$$
$$W_t|S_t = 0 \sim \log\mathcal{N}(\mu_w, \sigma_w^2)$$
\( R_t \): indicator of drive event
\( V_t \): velocity of ball handler towards hoop
\( Y_t \): distance between hoop and ball handler

\[
\begin{align*}
\frac{1}{V_t^{+}} | R_t = 1 & \sim \exp(\lambda_v) \\
Y_t | R_t = 1 & \sim \exp(\lambda_y) \\
V_t | R_t = 0 & \sim \mathcal{N}(\mu_v, \sigma_v^2) \\
Y_t | R_t = 0 & \sim \text{Unif}(0, \theta_y)
\end{align*}
\]
Post-up

- $U_t$: indicator of post-up event
- $A_t$: distance between on-ball defender and ball handler
- $Y_t$: distance between hoop and ball handler
- $H_t$: speed of ball handler

\[
A_t|U_t = 1 \sim \exp(\lambda_a) \\
Y_t|U_t = 1 \sim \log\mathcal{N}(\mu_y, \sigma_y^2) \\
H_t|U_t = 1 \sim \exp(\lambda_h)
\]

\[
A_t|U_t = 0 \sim \log\mathcal{N}(\mu_a, \sigma_a^2) \\
Y_t|U_t = 0 \sim \text{Unif}(0, \theta) \\
H_t|U_t = 0 \sim \log\mathcal{N}(\mu_h, \sigma_h^2)
\]
Inference

\[ h = \text{hidden state of event indicator } (S_t, R_t, U_t) \]

\[ x, y, z, ... = \text{sequences of observed states} \]

- Initialize \( \hat{P}(h_0), \hat{P}(x|h), \hat{P}(y|h), ..., \) and \( \hat{P}(h'|h) \) randomly
- Until convergence
  - **E Step**: For each sequence \( x, y, z, ... \), compute \( \hat{P}(h_0|x, y, z, ...), \hat{P}(h_t, h_{t+1}|x, y, z, ...), \hat{P}(h'|h) \) using forward-backward algorithm
  - **M Step**: Update the model parameters \( \hat{P}(h_0), \hat{P}(x|h), \hat{P}(y|h), ..., \) and \( \hat{P}(h'|h) \) using MLE
- Compute most likely sequence of hidden states, \( h = (h_0, ..., h_T) \) using Viterbi algorithm
Results
Estimated Defense Assignments and Events

- Lines represent estimated defense assignments
- Ball screen and drive actions are captured in the sequence
### Table 1: Defense Assignment Accuracy Comparison

<table>
<thead>
<tr>
<th>Model</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closest Defender</td>
<td>0.7597</td>
</tr>
<tr>
<td>Fixed $\Gamma$ Model (Franks et al.)</td>
<td>0.9179</td>
</tr>
<tr>
<td>Player Attraction based Model</td>
<td>0.9541</td>
</tr>
</tbody>
</table>

### Table 2: Event Detection Accuracy

<table>
<thead>
<tr>
<th>Event</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ball Screen</td>
<td>0.868</td>
</tr>
<tr>
<td>Drive</td>
<td>0.953</td>
</tr>
<tr>
<td>Post-up</td>
<td>0.994</td>
</tr>
</tbody>
</table>
### Table 3: Ball Screen Detection

<table>
<thead>
<tr>
<th>Actual</th>
<th>Prediction</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Positive</td>
<td>79</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>5</td>
</tr>
<tr>
<td>Negative</td>
<td>Positive</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>106</td>
</tr>
</tbody>
</table>

### Table 4: Drive Detection

<table>
<thead>
<tr>
<th>Actual</th>
<th>Prediction</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Positive</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>2</td>
</tr>
<tr>
<td>Negative</td>
<td>Positive</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Negative</td>
<td>58</td>
</tr>
<tr>
<td>Actual</td>
<td>Prediction</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Positive</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Negative</td>
<td>2</td>
<td>334</td>
</tr>
</tbody>
</table>
Heatmap for Selected Players

- Stephen Curry
- DeAndre Jordan
- LeBron James
Results for Golden State Warriors

- High $\gamma_0$: (S.C.) Stephen Curry, (K.T.) Klay Thompson
- Low $\gamma_0$: (A.B.) Andrew Bogut, (A.I.) Andre Iguodala
Results for Cleveland Cavaliers

- High $\gamma_0$: (K.I.) Kyrie Irving, (J.S.) J. R. Smith, (L.J.) LeBron James
- Low $\gamma_0$: (T.M.) Timofey Mozgov, (T.T.) Tristan Thompson
Questions?