State of Transition:

Estimating Real-Time Expected Possession Value in the NBA with a Spatiotemporal Transition Model and Player Tracking Data

Dan Cervone [1] Alex D'Amour [1]

[1]XY Hoops Group, Harvard University

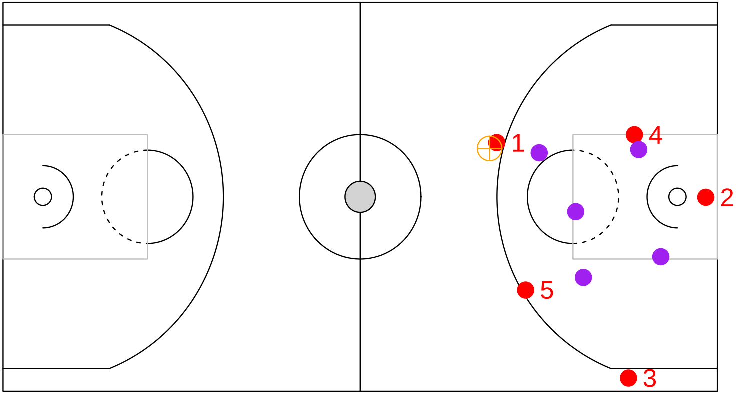
September 21, 2013

Optical tracking data

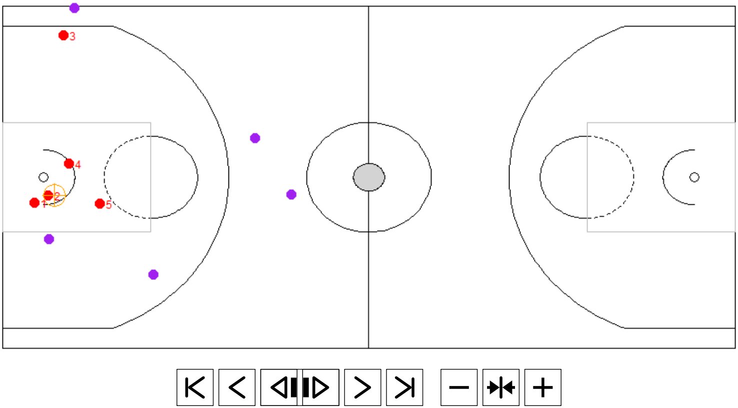
For the 2012-13 we have almost 800 million data points!

- 515 games
- 2D locations for all 10 players and 3 referees, 3D location for the ball
- 25 images per second
- Annotations (dribbles, passes, shots)

Optical tracking data



- Red points: Spurs (offense)
  - 1: Parker, 2: Jackson, 3: Green, 4: Duncan, 5: Diaw
- Purple points: Thunder (defense)
- Orange ⊕: ball


3
4
2
1
5

New frontiers for analyzing NBA offenses

#	W	B	#	W	B	#	W	B	#	W	B
1			7			13			19		
2			8			14			20		
3			9			15			21		
4			10			16			22		
5			11			17			23	Be7#	
6			12			18					

Table: Anderssen vs Kieseritsky, 1851

What can we learn about this chess game from "**23. Be7#**"?

New frontiers for analyzing NBA offenses

#	W	B	#	W	B	#	W	B	#	W	B
1	e4	e5	7	d3	Nh5	13	h5	Qg5	19	e5	Qxa1+
2	f4	exf4	8	Nh4	Qg5	14	Qf3	Ng8	20	Ke2	Na6
3	Bc4	Qh4+	9	Nf5	c6	15	Bxf4	Qf6	21	Nxg7+	Kd8
4	Kf1	b5	10	g4	Nf6	16	Nc3	Bc5	22	Qf6+	Nxf6
5	Bxb5	Nf6	11	Rg1	cxb5	17	Nd5	Qxb2	23	Be7#	
6	Nf3	Qh6	12	h4	Qg6	18	Bd6	Bxg1			

Table: Anderssen vs Kieseritsky, 1851

What can we learn about this chess game from "**23. Be7#**"?

Like chess matches, NBA possessions are often won/lost before the ball does/doesn't swish through the net.

- A teammate eludes the defense and gets open in the paint.
- The ballcarrier skips an easy shot to pass to a heavily defended teammate.
- With no look at the basket and no easy passes, the ballcarrier dribbles to a different spot.

Expected Possession Value

How many points is a team expected to score given the spatial evolution of its possession up to time t?

$$\text{EPV} = E[X|\mathcal{F}_t]$$

- X = number of points scored on this possession (**unknown**).
- $\mathcal{F}_t$ = space-time information of the possession up to time t.

EPV tells us

- When and how value was created *during* the possession
- Who created the value
- Who made the best decisions to increase their team's expected points

EPV is all about *what happens next*

Let A be the outcome of the next decision made by a player at time t.

- We could see, for instance, $A = \text{pass}$, $A = \text{take a shot}$, or $A = \text{dribble to basket}$.

By laws of probability,

$$\begin{aligned} E[X|\mathcal{F}_t] &= E[X|\mathcal{F}_t, A = \text{pass}]P(A = \text{pass}|\mathcal{F}_t) \\ &+ E[X|\mathcal{F}_t, A = \text{shoot}]P(A = \text{shoot}|\mathcal{F}_t) \\ &+ E[X|\mathcal{F}_t, A = \text{dribble}]P(A = \text{dribble}|\mathcal{F}_t) \end{aligned}$$

EPV is a weighted average of future EPVs, where the weights are transition probabilities.

Calculating EPV: step 1

More generally, let $\mathcal{S}$ be a binning of the full-resolution data into discrete *states* or *events* (e.g. "The point guard has the ball at the top of the arc").

Let $s_t \in \mathcal{S}$ be the state the possession is in at time t. Then:

$$\textbf{EPV: } E[X|\mathcal{F}_t] = \sum_{s \in \mathcal{S}} E[X|s_{t+\epsilon} = s, \mathcal{F}_t] P(s_{t+\epsilon} = s|\mathcal{F}_t)$$

Calculating EPV: step 1

More generally, let $\mathcal{S}$ be a binning of the full-resolution data into discrete *states* or *events* (e.g. "The point guard has the ball at the top of the arc").

Let $s_t \in \mathcal{S}$ be the state the possession is in at time t. Then:

$$\textbf{EPV: } E[X|\mathcal{F}_t] = \sum_{s \in \mathcal{S}} E[X|s_{t+\epsilon} = s, \mathcal{F}_t] P(s_{t+\epsilon} = s|\mathcal{F}_t)$$

Step 1 of calculating EPV is to define $\mathcal{S}$. Ideally,

- $E[X|s_{t+\epsilon} = s, \mathcal{F}_t] \approx E[X|s_{t+\epsilon} = s]$ for all $s \in \mathcal{S}$.
- $E[X|s_{t+\epsilon} = s]$ easy to calculate for all $s \in \mathcal{S}$ (ie, empirical average).

Calculating EPV: step 2

We need $P(s_{t+\epsilon} = s|\mathcal{F}_t)$ for all $s \in \mathcal{S}$.

- We have chosen $\mathcal{S}$ such that for all s, a good estimate of $P(s_{t+\epsilon} = s|\mathcal{F}_t)$ only needs:
  - $P(\text{shot in } (t, t+\epsilon)|\mathcal{F}_t)$
  - $P(\text{pass in } (t, t+\epsilon)|\mathcal{F}_t)$
- 6 different types of pass/shot events indexed by j.
  - Shots made, shots missed
  - Passes to each of 4 teammates
- Points in space-time corresponding to event j when player i has the ball follow an inhomogenous Poisson Process:

$$Y_j^i \sim PP(\lambda_j^i(t))$$

- All Y_j^i assumed independent.

Additional model details

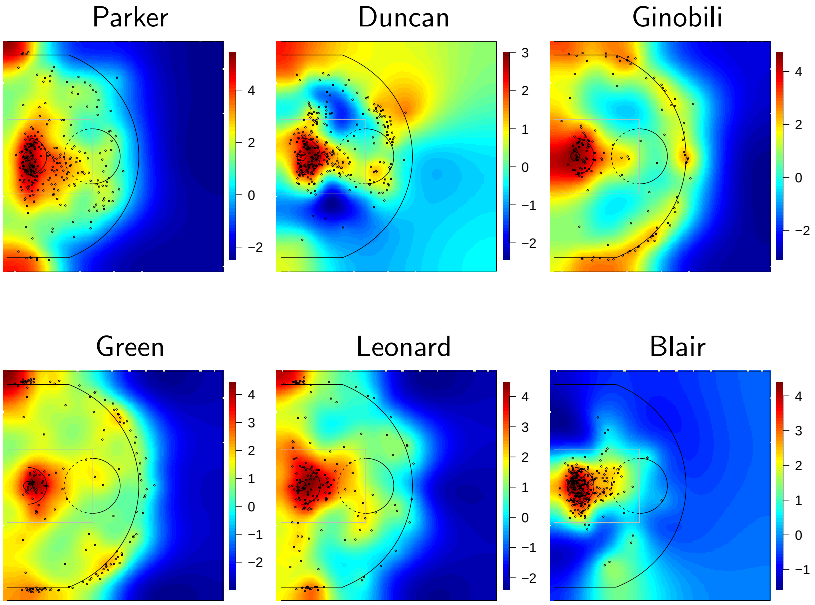
We additionally assume (i superscripts omitted):

$$\log(\lambda_j(t)) = \beta_j' W_j(t) + H_j(\zeta_t)$$

where

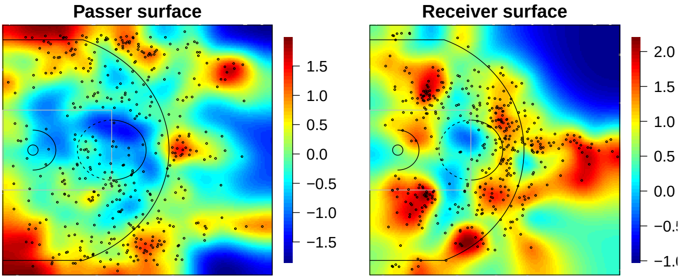
- $W_j(t)$ a p-vector of (possibly) time-varying covariates.
  - Distance to nearest defender; player i's velocity; has started dribbling, etc
- $\beta_j \in \mathbb{R}^p$ are coefficients for main effects.
- ζ_t is player i's location at time t.
- H_j is a spatial random effects surface (Gaussian Process).

Spatial random effect surfaces for *made shot events*

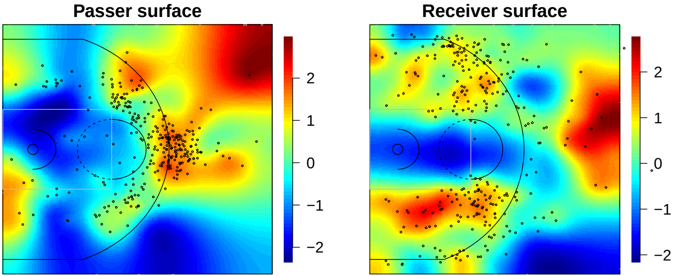


Spatial random effect surfaces for *pass events*

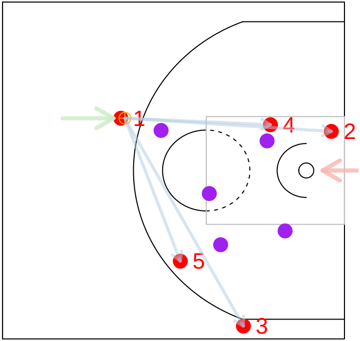
Parker to Duncan



Duncan to Parker



Putting it all together



- 1: Parker
- 2: Jackson
- 3: Green
- 4: Duncan
- 5: Diaw

Pass next:

$$\begin{aligned} E[X|\text{pass}] &= (0.78)(0.02) \\ &+ (1.08)(0.14) \\ &+ (0.84)(0.37) \\ &+ (0.85)(0.46) \\ &= 0.87 \end{aligned}$$

$P(\text{pass}) = 0.97$

Shoot next:

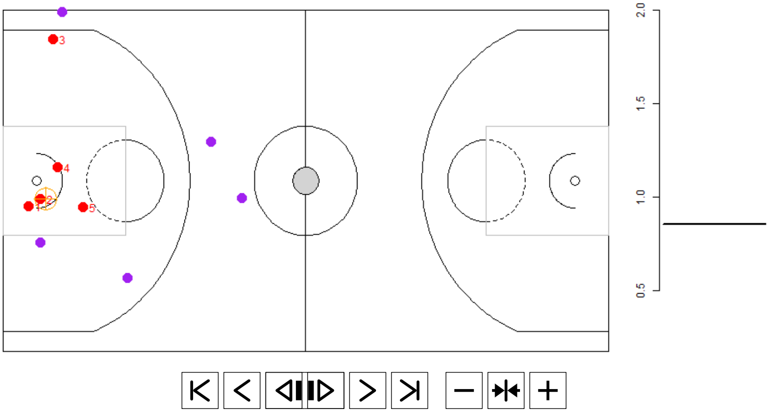
$$\begin{aligned} E[X|\text{shot}] &= (3.00)(0.18) \\ &+ (0.18)(0.82) \\ &= 0.69 \end{aligned}$$

$P(\text{shot}) = 0.03$

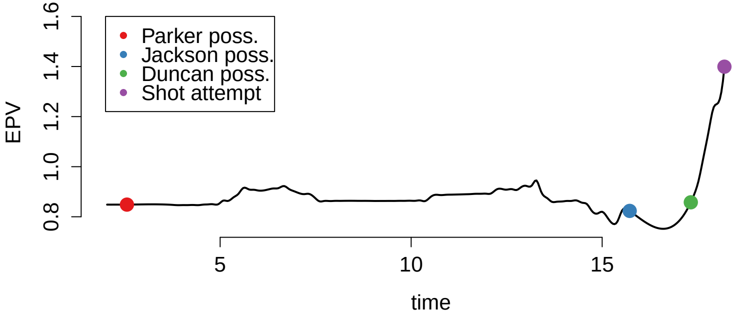
EPV:

$$\begin{aligned} (0.87)(0.97) &+ (0.69)(0.03) \\ &= \mathbf{0.86} \end{aligned}$$

EPV in real-time



EPV during a possession



EPV during game

Player	Summed ($\triangle$EPV)
Duncan	1.54
Jackson	0.50
Parker	1.44
Diaw	-0.02
Neal	0.40
Green	0.55
Blair	-1.21
Leonard	0.36

Table: Total change in Spurs' players' EPV while they were handling the ball during November 1, 2012 game against OKC

- Interpretation: Value added by players' decision-making in this game *relative to their usual value*.

The future of EPV

EPV is a powerful new tool for analysing NBA offenses:

- Diagrams where and how points are scored
- Can track EPV in real-time as the possession evolves
- Allows evaluation and quantification of players' decision-making
- In modeling EPV, we discover factors (including spatial effects) that influence players' decision-making

Nuances:

- Players can't systematically increase EPV
- EPV as a measure of skill

Future challenges:

- Incorporating defense
- Information-sharing across similar players (hierarchical models)