# NBA Court Realty

Dan Cervone, Luke Bornn and Kirk Goldsberry

New York Unversity, Simon Fraser University, and University of Texas (Austin)

## 1   Intro: the Basketball Court is a Real Estate Market

Continuously throughout every basketball possession, players control different regions of the basketball court. Some regions are more valuable than others, and players' control (or lack thereof) of valuable court space dictates the flow and strategy of a basketball possession for both the offense and defense. As professionals engaged in high-stakes competition, we assume players are rational actors, and the exchanges they make to acquire new court space represent a winning strategy for their team. This simple assumption allows us to infer the value (price) of court real estate based on player and ball movement alone.

   In this paper, we provide a definition of court space ownership, and infer the value of court space using SportVu player location data. Spatial tracking data has enabled a recent paradigm shift in basketball analytics [4, 7, 1, 3]; we add to this growing literature by focusing on *space*, itself, as the object of investigation. By modeling court space and court ownership, we reveal different valuable regions of the court among the NBA's players and teams, and insightful new metrics for both offense and defense. For instance, we can measure ballcarriers' off-ball impact on offense by calculating the value of the space freed up for their teammates to control. For analyzing defense, we can quantify how effectively different teams (and players) contain the offense (and particular players) within low-value regions of the court.

## 2   Quantifying Court Ownership

Unlike traditional property investments, players do not "own" any court space in any objective sense. Heuristically, though, it is easy to imagine a player completely owning the point he is standing on, having no control of a point very far away, and having moderate control of a point nearby (say 5 feet away)—except perhaps if another player is even closer to that point. With this in mind, we introduce a "weighted Voronoi" concept for defining each player's court space ownership at any point in time.

   We first divide the half-court into $M = 576$ equally sized cells, approximately $2' \times 2'$ each. For player $i$, $X^i(t)$ is a $M$-vector representing his investment (level of ownership) in each of the $M$ court cells at time $t$. The $m$th entry of $X^i(t)$ is inversely proportional to the distance between player $i$ and court cell $m$ at time $t$ only if no other player is closer to court cell $m$:

$$w_m^i(t) = \text{dist}(\text{player } i, \text{cell } m) \text{ at time } t \tag{1}$$

$$X_m^i(t) = \begin{cases} \frac{1}{1+w_m^i(t)} & i = \text{argmin}_j w_m^j(t) \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

ticketmaster®

Thus, court real estate is partitioned among the players according to the Voronoi diagram of player locations, and within each segment they control, players' investment in court space is inversely proportional to their distance from this space. Figure 1 illustrates this with a sample of our data.

Voronoi diagrams have previously been used for modeling sports data [9, 2], and offer several advantages. In particular, an offensive player's court ownership implicitly encodes information on the defensive positioning. When a defender closes on player $i$, player $i$'s Voronoi partition decreases, meaning more entries in the court ownership vector $X^i(t)$ are zero. Thus, regardless of the inferred court location prices, player $i$'s total real estate investment value decreases as he gets less open. Moreover, because of the weighting we use, this decrease is more dramatic when the defender is close, and negligible when a defender is far away, where even after approaching player $i$ he is still open.
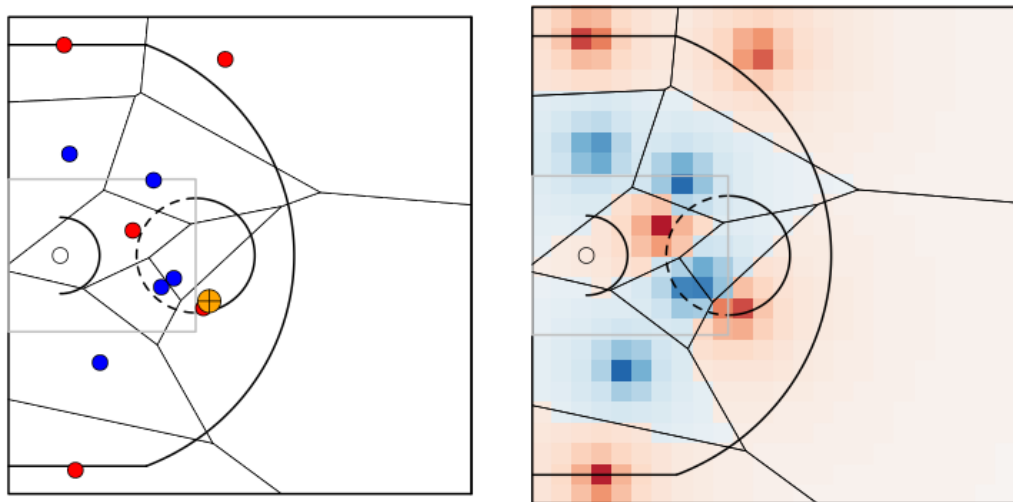


Figure 1: Example court space ownership map. A player's control of each court cell in his Voronoi segment (left) is inversely proportional to his distance from that cell (right).

## 3   Inferring Property Value

Like pricing in other financial markets, the value of NBA court property reflects a medium of exchange. Higher value regions can be identified if, based on players' movement and actions, they seem preferred to other regions on the court. For instance, when a player controlling a huge chunk of backcourt space passes to a teammate controlling a tiny section in the paint, this suggests that court space in the paint is more valuable than in the backcourt. The idea of inferring value based solely on asset transactions is used by Romer [8] to infer the relative value of picks in the NFL draft; our conceptual approach to valuing NBA court realty is very similar.

Passes between players offer clean, easily interpretable transactions of court real estate. However, unlike transactions in other markets, we don't expect the exchanges in court space implied by passes to be fair—because players in the offense cooperate instead of compete, passes should benefit the offense as a whole. Assuming players are generally rational decision-makers, when player $i$ passes to player $j$, this suggests (disregarding player-specific effects) player $j$ is in a more valuable position than player $i$; the team benefits from investing (ball control) in player $i$'s space instead player $j$'s space.

To formalize property value inference, let $\beta$ be a $M$-vector, with $\beta_m$ the price/value of court cell $m$. Thus, the total value of player $i$'s court real estate—his portfolio value—at time $t$ can be written

$$V^i(t) = [X^i(t)]'\beta$$

To estimate $\beta$ given player position data and pass events, we maximize:

$$L_\lambda(\beta) = \left[ \sum_{i,j} \sum_{t:\text{pass } i \to j} V^j(t) - \log \Big( \exp\big(V^i(t)\big) + \exp\big(V^j(t)\big) \Big) \right] - \frac{1}{2}\lambda||\beta||_2^2 \qquad (3)$$

$$\text{subject to } \beta_m \geq 0, \ m = 1, \ldots, M$$

where $i, j$ index players in the data, and $\lambda$ penalizes the $\ell_2$ norm of $\beta$. The positivity constraints on $\beta$, as well as the $\ell_2$ penalty, are not strictly necessary, yet they lead to more stable and interpretable performance of this model. Without the $\ell_2$ penalty, the objective function $L_\lambda(\beta)$ is constant for constant shifts in $\beta$, $\beta + c$; thus, the constraints on $\beta$ in (3) ensure that the lowest values for court space cells are near 0.

Our model (3) can be represented as a penalized logistic regression problem (or as a penalized Plackett-Luce model [6]), and easily fit using the `glmnet` package in R or similar software. Viewed as a logistic regression problem, the binary outcome is whether player $i$ passes to player $j$, or vice versa. Thus, our estimation of $\beta$ guarantees that the most probable feasible passes are those with the greatest gain in court property value from the passer to the pass target.

We choose $\lambda$ using cross-validation [5], which can also be used to evaluate the fit and predictive performance of our court space pricing model. Given the players involved in a pass, we predict the direction of the pass ($i \to j$ versus $j \to i$) extremely well—we are sometimes over 99% sure of the pass direction, yet measured on out-of sample test data, we are not overfitting.

## 3.1 Team-Specific Property Values

If we fit (3) using passing data from each team separately, we estimate team-specific court value surfaces $\beta^k$, where $k$ indexes the 30 teams in the NBA. Plotting $\beta^k$ thus reveals how teams value different areas of the court differently. Figure 2 shows the league average $\beta$, as well as the differences from the league average for three teams: Golden State, Houston, and the New York Knicks.

The league-wide $\beta$ plot reveals that space is most valuable near the basket, and in the corner 3 areas. There is a significant drop in value beyond 15 feet from the basket, before increasing again near the three point line (except in the middle of the court, where the value is low beyond the arc). Golden State's court value map is not drastically different from the league average (remember that our estimation of court space value does not use any shooting information), though there is more value in the wing three areas. Houston and New York both under-value and over-value mid-range shots, respectively, though Houston strongly values the area just inside the corner 3 (though this could mean they particularly value their corner 3 players being wide open, as an unguarded player in the corner would control significant space in front of him, as well).

## 3.2 Player-Specific Property Values

It is also possible to estimate player-specific property values, analogous to the team-specific parameters $\beta^k$. To do this, we modify our objective function (3) to include player adjustments to the court
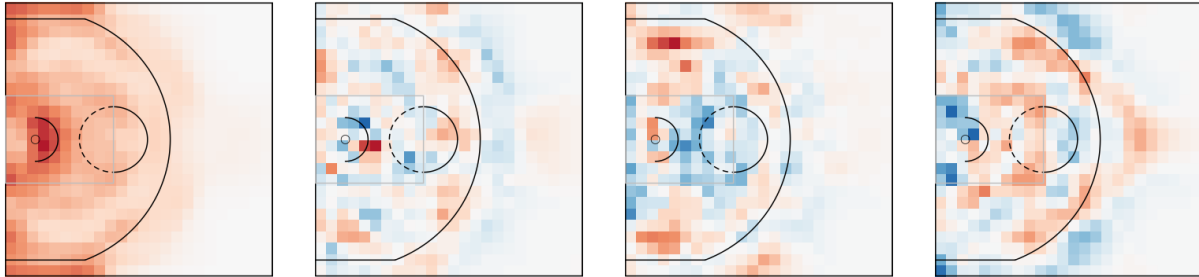
Figure 2: Court space value ($\beta$) plots for the 2014-15 NBA (left); and the difference ($\beta^k$ - $\beta^{\mathrm{NBA}}$) from this surface for Golden State, Houston, and New York, from left to right.

space value vectors $\beta$. Specifically, letting

$$W_i(t) = [X^i(t)](\beta + \alpha^i),$$

where $i = 1, \ldots, P$ indexes players, we maximize

$$L_{\lambda_1, \lambda_2}(\beta, \alpha^1, \ldots, \alpha^K) = \left[ \sum_{i,j} \sum_{t:\text{pass } i \to j} W^j(t) - \log \left( \exp\left(W^i(t)\right) + \exp\left(W^j(t)\right) \right) \right]$$
$$- \frac{1}{2}\lambda_1 ||\beta||_2^2 - \frac{1}{2}\lambda_2 \sum_i ||\alpha^i||_2^2 \qquad (4)$$

$$\text{subject to } \beta_k \geq 0, \ k = 1, \ldots, M.$$

This remains equivalent to a penalized logistic regression model, except note that we apply a different penalty to the player-specific deviations from $\beta$ ($\alpha^i$) than we do to $\beta$ itself. This makes sense—since $\alpha^i$ represent offsets from $\beta$, they should be closer to 0 than $\beta$ is. As with (3), we learn $\lambda_1$ and $\lambda_2$ using cross-validation, and fit this model (4) using data from each team separately (giving us a team-specific $\beta^k$ instead of a generic $\beta$).
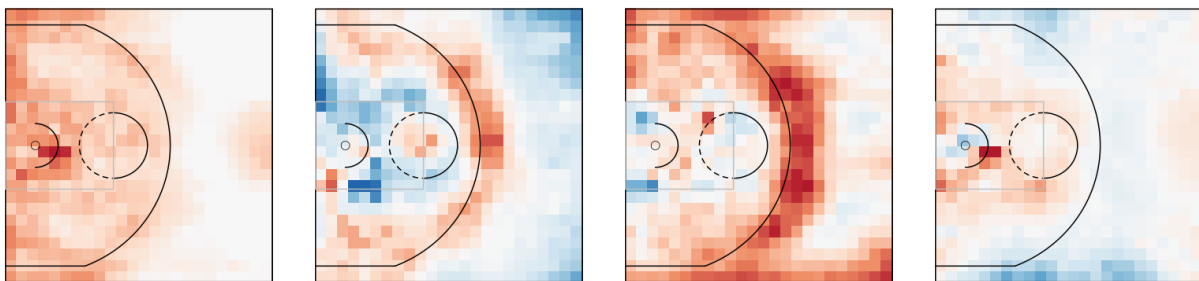


Figure 3: The leftmost plot is the 2014-15 Golden State Warriors' baseline property value plot ($\beta^k$); shown beside (from left to right) it are the player-specific court space value ($\alpha^i$) plots for Stephen Curry, Klay Thompson, and David Lee.

Figure 3 shows the $\alpha^i$ values for the 3 players on the Golden State Warriors, as estimated from their championship 2014-15 season. Relative to the rest of the team, Steph Curry and Klay Thompson both

have much higher court values around the three point line (particularly so for Thompson); David Lee's court space, on the other hand, is more valuable closer to the basket.

# 4   Property Value Derived Metrics

With estimates of court space prices, we can track the value of each player's real estate portfolio throughout any basketball possession. Figure 4 provides a glimpse of this, showing the court space each player controls and its associated portfolio value, at several moments during a possession from our data. Doing so provides useful quantifications of player positioning and spacing at the same level of temporal resolution as the original data, allowing basketball analysts to correlate these metrics with other events of interest. Two brief examples are presented below.
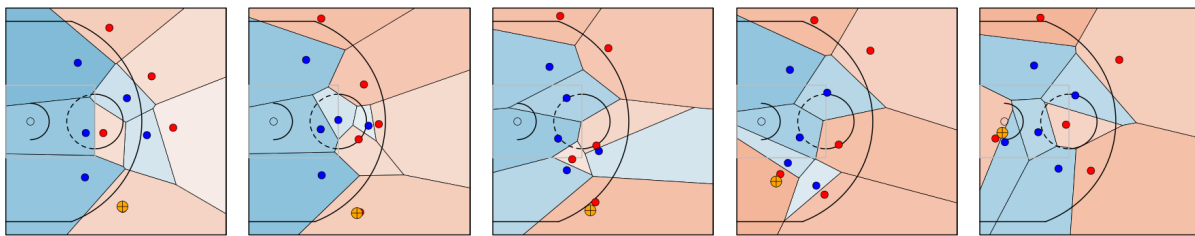


Figure 4: Example possession with each player's court ownership and real estate portfolio value illustrated. Offensive players are red and defensive players blue (darker color represents higher property portfolio value).

## 4.1   Ball Movement and Floor Spacing

Our court real estate model helps us quantify a player's effect on ball movement and floor spacing, two offensive features that—when executed well—ensure that players on the offense have higher total court real estate portfolio values. When this happens, either players are occupying more valuable regions, or simply being more open from the defense. To do this, we calculate the total portfolio of all players on the court when a particular player is in the lineup, versus when that player is removed. The top 10 and bottom 10 players by this metric are presented in Table 4.1.

## 4.2   Defensive Suppression

Though court space value is inferred only using offensive players' court ownership (during passes), our overall framework reveals valuable inferences about defensive spatial strategy. The weighted Voronoi court ownership definition implicitly values good defense, since when in valuable regions of the court, a player's portfolio value will drop when he is closely defended.

Expanding on this idea, we can calculate the average portfolio of an offensive player against each defensive team he faces, both while in control of the ball and not. Low portfolio values suggest that the player is either closely guarded when in valuable space, or mainly occupying low-value space—both of which suggest effective space suppression by the defense. In table 4.2, we present LeBron James'

| | Name | On-Court Team Portfolio Value | Off-Court Team Portfolio Value |
|---|---|---|---|
| 1 | Jae Crowder | 10.90 | 9.70 |
| 2 | Jordan Clarkson | 8.39 | 7.35 |
| 3 | Ryan Kelly | 8.45 | 7.42 |
| 4 | Marcus Smart | 10.53 | 9.78 |
| 5 | Brandon Bass | 10.50 | 9.75 |
| 6 | Wayne Ellington | 8.13 | 7.44 |
| 7 | Jonas Valanciunas | 10.35 | 9.71 |
| 8 | Evan Turner | 10.36 | 9.79 |
| 9 | Josh Smith | 11.42 | 10.89 |
| 10 | LaMarcus Aldridge | 9.80 | 9.31 |
| ⋮ | ⋮ | ⋮ | ⋮ |
| 271 | Kelly Olynyk | 9.82 | 10.27 |
| 273 | Lou Williams | 9.82 | 10.29 |
| 275 | Patrick Patterson | 9.84 | 10.31 |
| 274 | Jordan Hill | 7.44 | 7.97 |
| 275 | Ed Davis | 7.40 | 7.99 |
| 276 | Jared Sullinger | 9.65 | 10.43 |
| 277 | Kobe Bryant | 6.63 | 8.25 |
| 278 | Ronnie Price | 6.38 | 8.21 |
| 279 | Jeff Green | 8.38 | 10.77 |
| 280 | Rajon Rondo | 8.02 | 10.56 |

Table 1: On-court and off-court average team real estate portfolio values. Top 10 and Bottom 10 on-court − off-court differentials for 2014-15 are shown.

average portfolio value while on-ball and off-ball, taken separately against each opponent. Lower portfolio value scores represent the opposing defense's ability to contain LeBron in low value court space situations.

# 5   Conclusion

This paper combines economic reasoning and large-scale spatial data modeling in a novel analysis of NBA team and player strategy. The spatial impact of players, or specific offensive/defensive schemes that emphasize controlling particular regions of the court, is an important basketball analytics problem that has eluded quantification. By inferring the value of NBA court real estate, we enable new metrics for spacing and positioning, and uncover new axes to measure variation in team strategy.

# 6   Acknowledgements

| | Team | On-ball Portfolio Value | On-ball Portfolio Value |
|---|------|------------------------|-------------------------|
| 1 | Por | 2.02 | 2.09 |
| 2 | Was | 1.96 | 2.22 |
| 3 | NO | 1.95 | 2.10 |
| 4 | Atl | 1.95 | 2.03 |
| 5 | Den | 1.95 | 2.11 |
| 6 | LAC | 1.92 | 2.00 |
| 7 | Bos | 1.92 | 2.04 |
| 8 | OKC | 1.91 | 1.94 |
| 9 | Mem | 1.90 | 2.04 |
| 10 | GS | 1.89 | 1.86 |
| ⋮ | | ⋮ | ⋮ |
| 21 | Ind | 1.76 | 1.86 |
| 22 | Hou | 1.76 | 1.81 |
| 23 | Dal | 1.73 | 1.86 |
| 24 | Phi | 1.73 | 1.96 |
| 25 | Orl | 1.69 | 2.05 |
| 26 | Mil | 1.69 | 1.97 |
| 27 | LAL | 1.69 | 1.85 |
| 28 | Det | 1.67 | 1.88 |
| 29 | SA | 1.62 | 1.76 |
| 30 | Pho | 1.61 | 1.98 |

Table 2: On-ball and off-ball portfolio values for LeBron James in 2014-15, by opponent. Lower values indicate more suppression.

## References

[1] Dan Cervone, Alexander D'Amour, Luke Bornn, and Kirk Goldsberry. Pointwise: predicting points and valuing decisions in real time with nba optical tracking data. *Proceedings MIT Sloan Sports Analytics*, 2014.

[2] Sofia Fonseca, João Milho, Bruno Travassos, and Duarte Araújo. Spatial dynamics of team sports exposed by voronoi diagrams. *Human movement science*, 31(6):1652–1659, 2012.

[3] Alexander Franks, Andrew Miller, Luke Bornn, and Kirk Goldsberry. Counterpoints: Advanced defensive metrics for nba basketball. MIT Sloan Sports Analytics Conference. Boston, MA, 2015.

[4] Kirk Goldsberry. Courtvision: New visual and spatial analytics for the nba. In *2012 MIT Sloan Sports Analytics Conference*, 2012.

[5] Gene H Golub, Michael Heath, and Grace Wahba. Generalized cross-validation as a method for choosing a good ridge parameter. *Technometrics*, 21(2):215–223, 1979.

[6] John Guiver and Edward Snelson. Bayesian inference for plackett-luce ranking models. In *proceedings of the 26th annual international conference on machine learning*, pages 377–384. ACM, 2009.

[7] Rajiv Maheswaran, Yu-Han Chang, Aaron Henehan, and Samantha Danesis. Deconstructing the rebound with optical tracking data. In *The MIT Sloan Sports Analytics Conference. Boston*, 2012.

[8] David Romer. Do firms maximize? evidence from professional football. *Journal of Political Economy*, 114(2):340–365, 2006.

[9] Stefan Schiffer, Alexander Ferrein, and Gerhard Lakemeyer. Qualitative world models for soccer robots. In *Qualitative Constraint Calculi, Workshop at KI*, volume 2006, pages 3–14, 2006.