



The Three Dimensions of Rebounding

Rajiv Maheswaran, Yu-Han Chang, Jeff Su, Sheldon Kwok, Tal Levy, Adam Wexler, Noel Hollingsworth

Second Spectrum, Inc.

Los Angeles, CA

rajiv@secondspectrum.com

Abstract

The recent spread of tracking technology in sports is bringing about a new era in analytics where we can deconstruct things we previously understood as one thing. We consider rebounding in basketball. Until recently we would get at most one piece of information after a missed shot: the name of a player that got the rebound. In this paper, we (1) describe the full timeline of a rebound, (2) develop metrics for the various dimensions of this timeline using novel techniques and (3) apply them to calculate individual player abilities in these dimensions.

1 Introduction

The danger in creating a statistic is that it combines many different dimensions of performance into one thing. One reason is that we simply have not had the necessary data. The recent spread of tracking technology in sports is bringing about a new era in analytics where we can deconstruct metrics we previously understood as one skill into the many properties we understand them to be when we watch the game.

We consider rebounding in basketball. Until recently we would get at most one piece of information after a missed shot: the name of a player that got the rebound. Getting the ball back after a miss is important, so we count up these rebounds, advance the notion of “rebounding” as a skill and try to understand it. But when watching the game, we know that all rebounds are not the same. One could get the ball by (1) being the only player within 10 feet of the ball, (2) have it bounce to you after going through the hands of three others who weren’t able to corral it or (3) fight off two players and jump over a third to get it. These aren’t the same, though they are all marked as a rebound. We are aware that a rebound has a timeline, a life cycle, from the release of a shot to the time it is controlled, filled with actions and movements by 10 players. Until now, that data was invisible.

Rebounding has been a ripe area for analytics in tracking data. Early work showed relationships between shot location and reboundability, and trajectory analysis of the ball after misses [1]. This was followed by work that investigated strategies for crashing the boards [2]. These initial forays have dabbled into the possibilities, but have not fully specified the rebounding process. In this paper, we (1) describe the full timeline of a rebound, (2) develop metrics for the various dimensions of this timeline using novel techniques and (3) apply them to calculate individual player abilities in these dimensions.

2 The Rebounding Timeline

We consider rebounds from missed field goal attempts in the NBA. The traditional measure of rebounding was the number of rebounds a player would get in a game. Advanced measures would address that players might have different numbers of opportunities in a game, and would normalize this by minutes or possessions. Here, we describe the full timeline of a rebound and address the various dimensions involved.

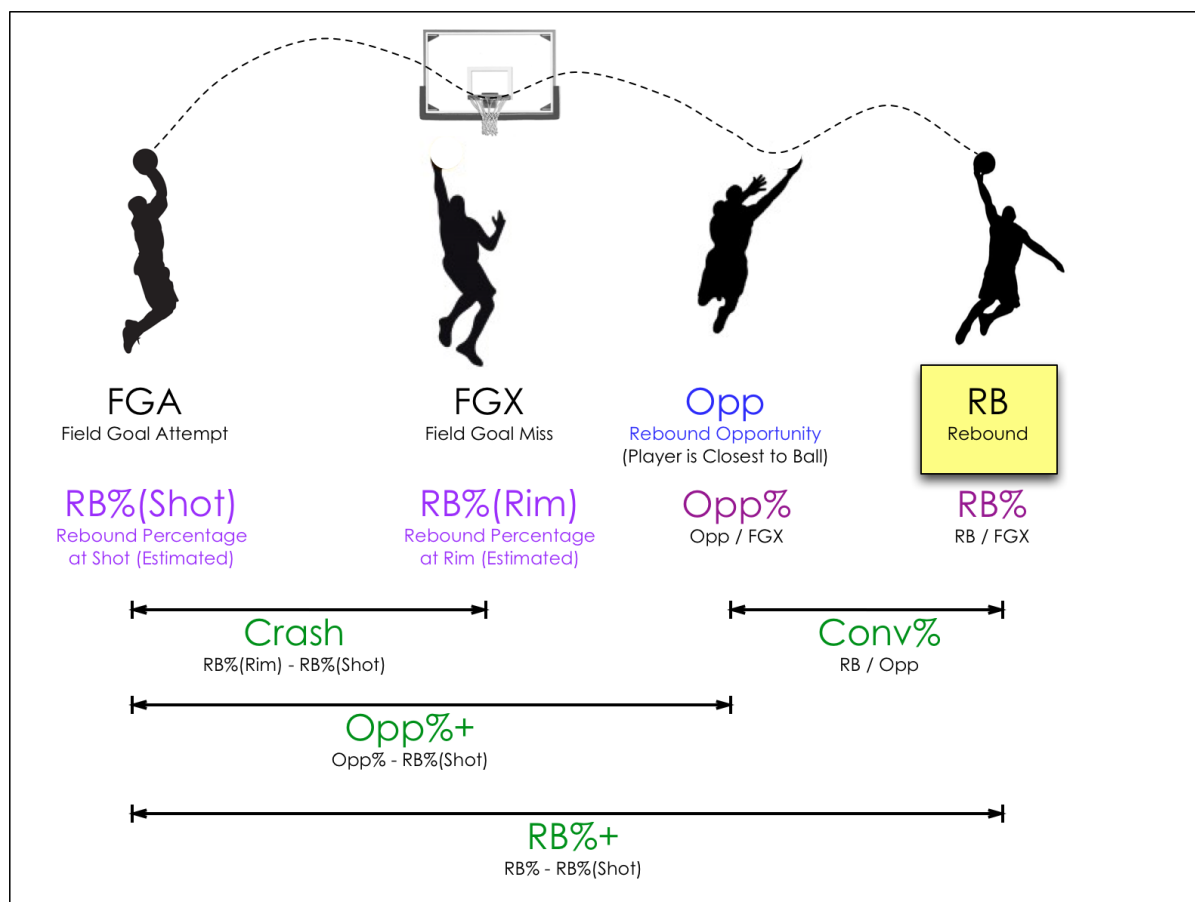
We begin with a field goal attempt. We understand by watching the game that at this point some players are immediately in a better position to get a rebound than others. The player standing closest to the basket is far more likely to get the rebound than someone standing outside the three-point line. This is the first dimension of rebounding: **Positioning**. Positional rebounders are those who by skill, luck or design are in a good place to get the rebound immediately at the shot attempt. Once the ball is released, players crash the boards and block others out to reposition themselves until the ball hits the rim: we denote this as the *Crash* phase. Once the ball hits the rim and bounces off, players have the opportunity to pursue the ball. Players who are closest to the ball after it leaves the rim have an *Opportunity* to get a rebound. The time from initial positioning at the field goal attempt and the opportunity to get the rebound is a phase where players can translate position to opportunity. The ability to do this is denoted as **Hustle**: the second dimension of rebounding. When a player has an opportunity to get a rebound, by



being the closest player to the ball, we observe that they are not always able to secure the rebound. The ability to turn opportunities into rebounds is the third dimension of rebounding: **Conversion**. The simple formula is:

$$\begin{aligned}\text{Rebounding} &= \text{Opportunities} \times \text{Conversion} \\ \text{Opportunities} &= \text{Positioning} + \text{Hustle} \\ \text{Rebounding} &= (\text{Positioning} + \text{Hustle}) \times \text{Conversion}\end{aligned}$$

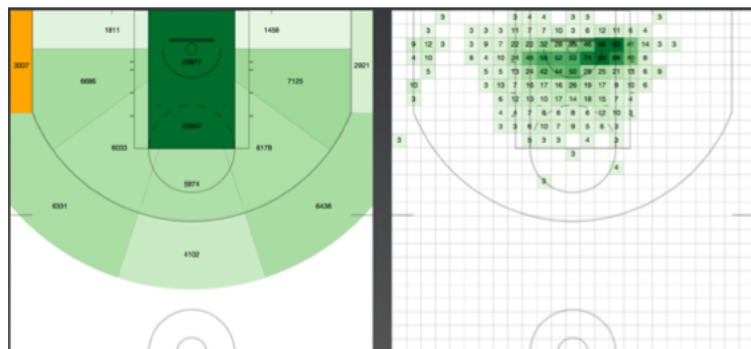
We will come to see that the breakdown is not quite as simple as this because the dimensions are not initially independent. In the following sections, we will take a look at the various aspects of rebounding discussed above, show how we can develop metrics to analyze performance in each of these areas and see how it applies to the NBA.



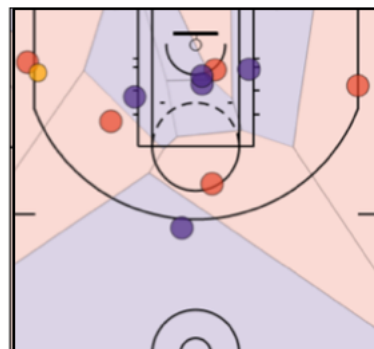
3 Positioning

We first address how to value the initial position of the players when the shot is taken. This is the most significant and difficult of the metrics to establish. We are trying to give a value to the real estate that each player owns at the time of the shot. This breaks down into two questions: (1) what is the real estate for each player? (2) what is it worth?

To address the first question, we apply the technique of using Voronoi (or Dirichlet) tessellations [3,4]. Voronoi tessellations are often applied to problems involving spatial allocation. These tessellations partition a space into Voronoi cells given a number of points in that space. For any point, it is the intersection of the self-containing half-spaces defined by hyper-planes equidistant from that point to all other points. This is a fancy way of saying a player's cell is all the points on the court that are closer to the player than any other player. An example tessellation is shown in the figure below. If all players were equally capable they should be able to control any rebound that occurred in this cell. We understand that players are not equally capable however this establishment of real estate is to set a baseline for performance. Over performance or under performance of this baseline will be indicative of their ability.



Rebound Spatial Distribution Based on Shot Location



Voronoi Tessellation

To address the second question, we condition on where the shot was taken and calculate a spatial probability distribution of where all rebounds for similar shots were obtained. In the figure above, we show a spatial distribution of rebounds from shots missed in the left-corner-three area. This is only for illustration. For this paper, for each shot attempt, we choose a collection of shots closest to the shot location that provides enough samples to construct a distribution. This distribution captures the value of the real estate across the court for a given shot.

To assign each player a value for initial positioning, i.e., the value of the real estate at the time of the shot, we integrate the spatial distribution over the Voronoi cell for that player. This yields the likelihood of that player getting the rebound if no one moved when the shot was taken and they controlled their cell. We note that because we use the distribution of location of the rebound conditioned on the shot, it is not a matter of controlling more area or even necessarily area close to the basket, but the most valuable area for that shot. While the most valuable areas are almost always close to the basket, there are some directional effects. We applied this analysis to the NBA using tracking data from the STATS SportVU system implemented league-wide this year for players who were on the court for at least 500 field goal attempts that were missed.

In the tables below, we list the top ten players who have the highest initial positioning. The left two are for power forwards and centers and the right two are for point guards, shooting guards and small forwards. We also separate by defensive and offensive rebound percentages. If we look at the $DRB\%(Shot)$ for power forwards and centers, we see that Larry Sanders, Roy Hibbert and Meyers Leonard based on their position at the time of a shot have over a 30% chance of getting a defensive rebound according to the value of the real estate they have at that time. Similarly, on offense Reggie Evans is at an initial position at the time of the shot to get a rebound 15.8 percent of the time. On offense, one way to be in a good position to get a rebound is to be the shooter who takes shots close to the rim, but we see that the lists for offense are primarily not this type of player, with Zach Randolph and Carmelo Anthony being notable exceptions. The $DRB\%(Shot)$ for the bigs (PF, CTR) reflects players who one would typically associate with spending a lot of time at the rim during a possession.

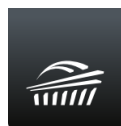
ORB%(Shot) [PF,CTR]	
Reggie Evans	15.8
Nikola Pekovic	14.6
Bismack Biyombo	14.6
Zach Randolph	14.5
Joel Freeland	14.0
Rudy Gobert	14.0
Jeff Pendergraph	13.6
DeAndre Jordan	13.3
Omer Asik	13.3
Jordan Hill	12.7

DRB%(Shot) [PF,CTR]	
Larry Sanders	32.0
Roy Hibbert	30.9
Meyers Leonard	30.6
Rudy Gobert	29.2
Ian Mahinmi	29.0
Greg Stiemsma	29.0
Andrew Bogut	29.0
Jeff Adrien	28.7
Omer Asik	28.7
Jermaine O'Neal	28.6

ORB%(Shot) [PG,SG,SF]	
Jeremy Evans	8.2
Chris Singleton	7.7
Luc Richard Mbah a Moute	7.5
Carmelo Anthony	6.8
Tyreke Evans	6.5
Tobias Harris	6.4
Shaun Livingston	6.1
P.J. Tucker	5.8
Will Bynum	5.8
Landry Fields	5.6

DRB%(Shot) [PG,SG,SF]	
Dante Cunningham	19.6
Shane Battier	19.0
Jeremy Evans	18.4
Marvin Williams	18.0
Tobias Harris	17.3
Omri Casspi	17.2
Carmelo Anthony	17.0
Chris Singleton	16.6
Luc Richard Mbah a Moute	16.6
Shawne Williams	16.6

The rebound percentage at the shot or initial position value is important because it is the launching pad for the dimensions that come after it.

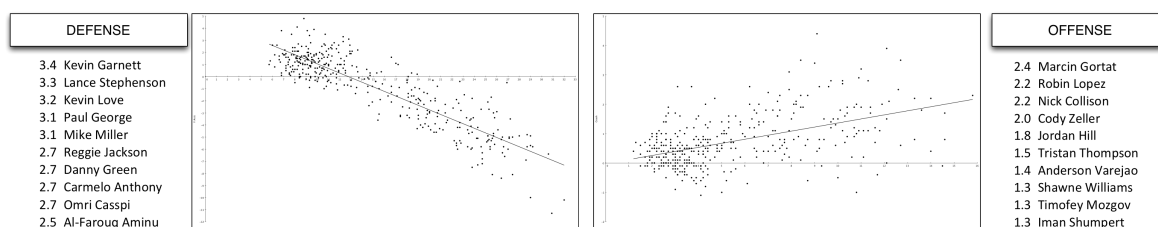


4 Crashing and Blocking Out

We first look at the Crash phase. To analyze this, we look at the trajectory of the ball and calculate the time that it gets closest to the center of the rim. At this point, we reapply the Voronoi-based analysis and calculate the rebound percentages of each player, i.e., the value of the real estate that each player has at the time the ball hits the rim. The change in this percentage from the time the shot is taken to the time it hits the rim is the value or likelihood the player had added during the phase. Players can add value by crashing the boards, i.e., moving closer to the basket towards places where the rebound is likely to go, or by blocking out, i.e., preventing other players by taking valuable real estate that is already established. A straightforward metric for the crash phase would be to subtract the rebound probability at the shot from the rebound probability at the rim. The issue is that the ability to add probability is not independent from the probability at the shot. Consider a case of a defensive player who plays close to the basket. They are occupying high value real estate and once the shot is taken, other players are going to start coming into this real estate. It is difficult for players with high initial positioning value to have positive crash deltas. Now consider a player out by the three-point line. Their initial value is very low and moving any significant distance toward the rim will give them a positive crash delta. Thus, it is not fair to compare these players on the same scale.

To address this, we look at the relationship of the raw crash deltas (the difference between the probability at rim and probability at shot) compared to the probability at shot. The relationships are shown in the figure below with defense on the left and offense on the right. The X-axis is probability at shot and the Y-axis is raw crash delta. The scatter plot has data points for each player. The line is a plot of a linear regression of these variables [5]. For defense, we see that it matches the example. Players with high initial positioning values have negative crash deltas and players with low initial positioning values have positive crash deltas. For offense, we see that all players have positive crash deltas. This is reasonable as offensive players are generally further away from the basket than defensive players and are able to gain probability after a shot. Interestingly the slope of the regression is positive indicating that players with higher initial probabilities make larger gains. This would indicate that players who are closer to the basket on offense tend to try for offensive rebounds more than those further away, though as the scatter plot shows, there is variation in behavior.

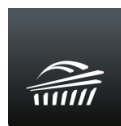
In order to normalize for this effect, we subtract the value of the regression at the player's initial positioning value from the raw crash delta to form the players Crash value. Intuitively, the value indicates how much more probability is added by this player beyond what a player with similar initial positioning would add. We will apply this normalization methodology to all the metrics we develop because as we discovered and will be shown, the initial positioning affects the other dimensions and we must control for it.



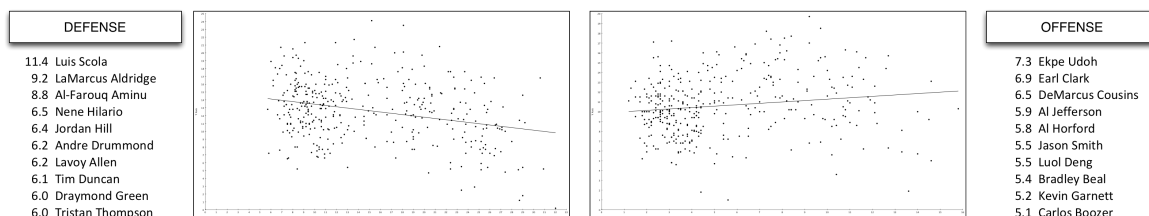
Looking at the list, we see the Kevin Garnett and Marcin Gortat are the best at blocking out / crashing at the defensive and offensive ends, respectively. Kevin Garnett adds 3.4% of rebound probability in the crash phase more than what would be expected given his initial position. The interesting part of the normalization is that it allows one to compare across positions. We see that the Pacers have two wings (Lance Stephenson and Paul George) that are incredibly good at crashing and blocking out on the defensive end.

5 Hustle

A player has an opportunity to rebound the ball if they are the closest player to the ball once the ball gets below 10 feet (or if they possess the ball while it is above 10 feet). The player with the first opportunity may not get the rebound so multiple opportunities could be created after a single field goal miss. We tallied the number of field goal misses for which a player generated an opportunity for themselves and divided by the number of field goals to create an opportunity percentage metric. This indicates the percentage of field goal misses for which that player ended up being closest to the ball at some point. The ability for a player to generate opportunities beyond his initial position is the second dimension of rebounding: **Hustle**.



While for small sample sizes, opportunities might reflect biases in ball direction (i.e., a player may not have chances at the ball and be penalized), for large sample sizes it indicates the player's ability to get to the ball, hence the name "Hustle". A straightforward Hustle metric would be to subtract the Positioning value from the Opportunity Percentage but as we noted from the Crash exercise, the ability to generate opportunities might be correlated with initial positioning. Thus, we performed the same regression and adjustment as we did in Crash case to come up with our Hustle metric. The regressions are shown in the figures below. We note that while not as strong as in the Crash metric, there are still some effects. Players with higher initial positioning have lower raw Hustle numbers as opposed to those with lower initial positioning on defense and vice versa on offense.



The lists on the left and right in the figure list the top ten players for Hustle on defense and offense, respectively. The numbers indicate how much additional probability these players add in terms of creating opportunities beyond initial positioning compared to players with the same initial positioning. Luis Scola, LaMarcus Aldridge and Al-Farouq Aminu are in a class by themselves on the defensive end while Ekpe Udoh, Earl Clark and DeMarcus Cousins lead Hustle on the offensive end. Again, we see that this methodology allows us to compare marginal value added across multiple positions.

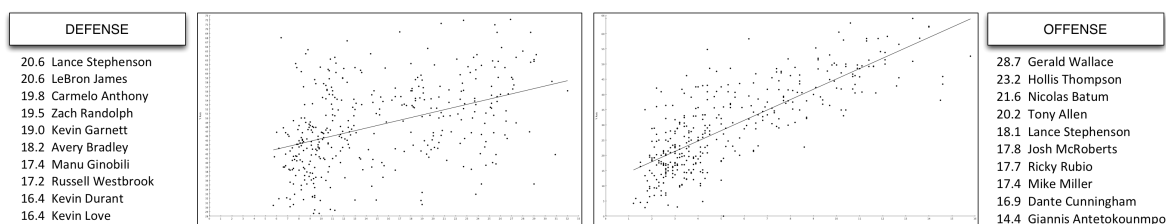
6 Conversion

The reason that there are often multiple opportunities for rebounds for every missed shot is that being closest to the ball does not mean that a player will convert it into a rebound. Thus, the third dimension of rebounding: **Conversion**. The raw conversion metric for players is calculated simply by dividing the number of rebounds obtained by the number of opportunities generated.

The top players for offense and defense for raw conversion rate are displayed in the table on the right. It is populated with players that are thought of as the best rebounders of each type in the league. Raw conversion rate seems to be what people primarily think of when they think of good rebounders and it matches intuition in that it represents players who gobble up the ball whenever they are close to it. These are the players who are least likely to have rebound taken away from them. We note DeAndre Jordan's presence on both lists.

DEFENSE	OFFENSE
71.9 Kevin Garnett	65.0 DeAndre Jordan
71.2 Zach Randolph	58.4 Andre Drummond
70.9 Kevin Love	57.9 Trevor Booker
69.3 Carmelo Anthony	56.5 Jordan Hill
69.2 Dwight Howard	54.8 Gerald Wallace
67.3 LeBron James	54.1 DeJuan Blair
65.8 DeAndre Jordan	53.8 Robin Lopez
65.6 Blake Griffin	51.7 Anderson Varejao
65.3 Lance Stephenson	51.4 Samuel Dalembert
65.0 Bismack Biyombo	51.3 Kenneth Faried

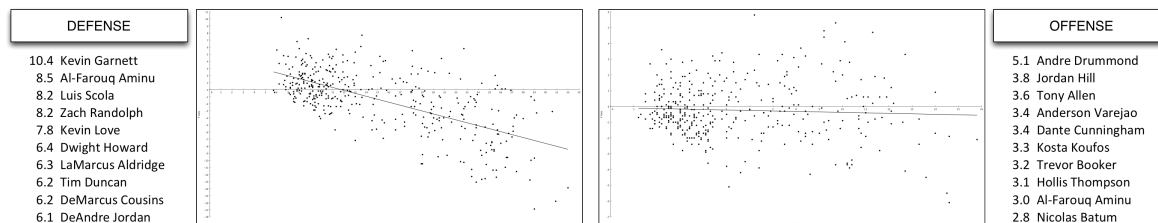
As for previous metrics, the ability to convert is correlated with players' initial positioning. We apply the same methodology and regress conversion percentage to initial position value. The scatter plots are displayed below. Interestingly we see that for both defense and offense, it is easier to convert the higher your initial positioning. This matches the intuition that opportunities for people with high initial positioning values are closer to the basket and in tighter cells so an opportunity means that the ball is closer to them. Those with lower initial position might have an opportunity where the ball is further from them and therefore more difficult to get. We make the regression adjustments to yield the top players for defense and offense below. For defense, the list is primarily the same but highlights three guards who perform significantly above expectation: Avery Bradley, Manu Ginobili and Russell Westbrook. On offense, the relationship between positioning and conversion is much stronger and we see that it is primarily guards and wings who are the outliers in terms of converting above expectation.





7 Hustle and Conversion

Here, we take a look the combined efforts of Hustle and Conversion. It is unclear if Positioning is a matter of skill, luck or design, but it is more clear that the other two dimensions indicate some type of ability. We look at which players have the best combination of these abilities by looking at players' rebound percentages, i.e., the number of rebounds obtained per field goal miss for which they are on the court and subtract the probability they have due to positioning. We again apply the regression modeling, which is displayed below. We see a strong effect in defensive rebounding, where it is difficult to improve when your initial positioning is high and thus adjust for that. Interestingly, the effect seems to be relatively flat on the offensive end.



Looking at the lists for offensive and defensive players, we again see a list that is highly reflective of players who are considered elite at each category. If one assumes that positioning is not a skill, then this metric would be the one that would most reflect an overall “rebounding” ability. We note that Al-Farouq Aminu is an elite rebounder at both ends of the court and the only player in the top ten of both lists (#2 defense, #9 offense). Expanding a little, Kevin Love (#5 defense, #19 offense), DeMarcus Cousins (#9 defense, #12 offense) and Kenneth Faried (#13 defense, #15 offense) emerge as players who are top 20 on both offense and defense. Another notable player is Iman Shumpert who is the only guard who appears highly in both lists (#23 defense, #31 offense,).

8 Conclusion

In this work, we have deconstructed rebounding into three dimensions: Positioning, Hustle and Conversion. Each captures a different skill that contributes to overall rebounding performance. By leveraging a Voronoi-tessellation approach with a spatial probability distribution, we are able to calculate new metrics for Positioning that serves as the basis for metrics for all the dimensions. Also, importantly, we leverage a regression-based normalization to account for the fact that the other dimensions are not initially independent of Positioning.

The high-level arc of this type of work is that with data and the right approach we are able to build analytics that deconstruct what were coarse metrics into skills that we know exist as we watch the game. This allows us to analyze something like rebounding in a manner that is more reflective of what is going on in the game. It can also highlight why particular players are good or not so good in this area of performance and give coaches and front offices indications of where to improve. In addition, the normalization allows one to find people who contribute across position. Both traditional and previous advanced metrics were heavily biased towards giving credit to bigs who spend a lot of time near the basket.

A key issue to be addressed in the future is the determination of whether Positioning is skill, luck or design. It may be the case that the answer is different for different players. One can envision that the decomposition of rebounding can get even more nuanced. One hypothesis is that players can anticipate that a shot is likely to be taken and thus, move into a position to rebound. This would indicate that the life cycle of a rebound would start even before the shot. One can also further deconstruct Hustle into Crash and post-Crash components.

The Three Dimensions of Rebounding represent initial forays for basketball into establishing baseline expectations for various aspects of the game and judging performance relative to that expectation as opposed to using absolute metrics. In this case, the baseline is Positioning. This approach has already permeated baseball but spatiotemporal sports such as basketball have much more complexity in establishing these expectation levels. With the NBA having league-wide adoption of player tracking, we anticipate that we are at a tipping point for the development of these types of models and thus, fundamentally changing the way the sport is analyzed.



References

- [1] Rajiv Maheswaran, Yu-Han Chang, Aaron Heneghan and Sam Danesis, "Deconstructing the Rebound with Optical Tracking Data." MIT Sloan Sports Analytics Conference, 2012.
- [2] Jenna Wiens, Guha Balakrishnan, Joel Brooks and John Guttag, "To Crash or Not to Crash: A quantitative look at the relationship between offensive rebounding and transition defense in the NBA" NBA. 2012 MIT Sloan Sports Analytics Conference.
- [3] Franz Aurenhammer (1991). Voronoi Diagrams – A Survey of a Fundamental Geometric Data Structure. ACM Computing Surveys, 23(3):345-405, 1991.
- [4] Atsuyuki Okabe, Barry Boots, Kokichi Sugihara & Sung Nok Chiu (2000). Spatial Tessellations – Concepts and Applications of Voronoi Diagrams. 2nd edition. John Wiley, 2000
- [5] Neter, John, William Wasserman, and Michael H. Kutner. Applied linear statistical models. Vol. 4. Chicago: Irwin, 1996.