ASID: A New Metric for Analyzing Spacing Impact in the NBA

Chris Spennato - 40162948

December 6, 2021

1 Introduction

Spacing has always been important in the National Basketball Association (NBA), but it has recently received an even bigger focus with the higher volume of three-point shooting in the league today. The increasing ability of players to shoot from beyond the arc forces defenses to play tighter on the perimeter, in turn creating more space for the offensive team to run plays and find separation from the defense.

Some work quantifying spacing exists, but remains limited. Frencken and Lemmink (2008) proposed two variables to analyse the flow in small-sided soccer games - the centroid position and surface area of a team. Frencken et al. (2011) further applied both variables in depth to soccer. Likewise, Bourbousson et al. (2010) applied the centroid position, along with a stretch index, to analyse team interaction in basketball. Franks et al. (2015) moreover employed surface area, through the use of convex hulls,[1] to quantify individual defensive performance in basketball, while Metulini et al. (2017) also used convex hulls to analyse players' movements, again in basketball.

Additionally, several sports analytics blogs apply the CHAD (Convex Hull of the Area of the Defense), an extension of the variables proposed by Frencken and Lemmink (2008), to visualise and quantify the area covered by a defense. See Squared Statistics[2] and The Spax[3] for examples.

However, none of this work aims to directly analyse a team's spacing, nor any potential effects on a team's spacing; instead using spacing to investigate other aspects of the game. In this paper, I propose a metric, ASID (Average Spacing Impact on the Defense), which will quantify a NBA team's effect on their opponent's defensive spacing. This metric allows us to directly analyse spacing and can provide a useful tool to examine a team's offensive scheme and gauge its effectiveness at stretching the defense.

[1] A convex hull is the smallest convex polygon that encloses all points in a set.
[2] https://squared2020.com/2015/11/08/building-nba-defenses-using-the-convex-hull/.
[3] https://www.thespax.com/nba/analyzing-nba-spacing-with-player-tracking-data/.

2 Data

I consider player tracking data and event data from a portion of the 2015-2016 NBA season. It contains 632 games from October 27, 2015 to January 23, 2016. The SportVU system collected this data and the NBA provided it publicly for a short period. Fortunately the data remains in archived form.[4] While it is not a full season of data, it will nonetheless be enough to calculate and present the metric usefully.

The dataset is separated by game and contains both tracking data and event data. The tracking data consists of (x, y) positions for all players on the court, as well as the (x, y, z) coordinates of the ball, captured at twenty-five frames per second. These positions are split into events, each labeled with a unique ID per game. The event data consists of all events occurring during a game, labeled with information such as event type, players involved, game clock, etc. Using the unique event IDs, I am able to join the positional tracking data with the event data. This allows me to properly attribute the tracking data to varying event types.

3 Methods

All my work derives from building on code provided by github user linouk23.[5] Their script for visualising a play provided a code structure upon which to expand. To calculate the average CHAD during a game, I extract only plays that contained a shot. These are any events in the dataset labeled as either a make, or a miss. I split the shots based on the attacking team. For each shot, I calculate the defense's average CHAD. I do this by calculating the CHAD at each frame up until the shot time, as recorded in the event data. This results in two average CHADs, one for the home team, and one for the away team. When calculating these CHADs, I only include plays that had at least 3 seconds of usable tracking data for the play.

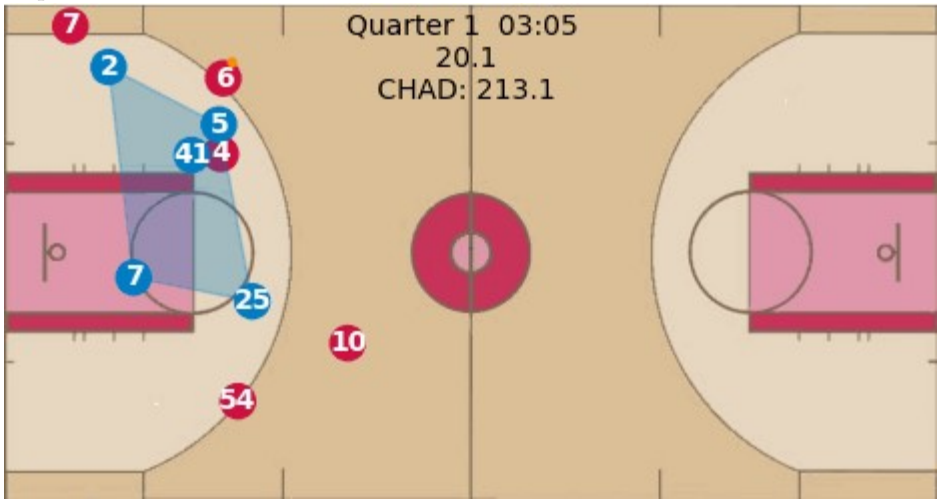
Figure 1 provides an example of a calculated CHAD, during a Toronto Raptors versus Dallas Mavericks game.

I then calculate the average CHAD for each team, for all games in the dataset. With this data, I calculate each team's average CHAD over the partial season contained in our dataset. When calculating this season average CHAD, I only include games in which at least half of the shots taken are in the average CHAD. This is to account for the fact that some plays are mislabeled or do not have adequate tracking data.

Using this new season average CHAD, I calculate the deviation from this average for each team, for each game. I attribute this deviation to the opposing team in each game, denoting it as the Spacing Impact. Because Spacing Impact is measured as a change in surface area, which is not a particularly useful measurement on its own, I convert it to percentage, referred to as Percent Spacing

[4] Dataset available here: https://github.com/sealneaward/nba-movement-data.
[5] Code available here: https://github.com/linouk23/NBA-Player-Movements.

Figure 1: Dallas Mavericks (in blue) CHAD during a play against the Toronto Raptors



Impact.

Finally, I calculate each team's average Percent Spacing Impact over the partial season. This is simply the average of their Percent Spacing Impact in each game. This gives my final metric, the Average Spacing Impact on the Defense, or ASID.

4 Results

In total 64.5% of shots in the dataset were usable, with 409 games used to calculate the season average CHADs. Teams thus have an unequal number of games. The average number of games used per team was twenty-seven. The team with the most games included were the Portland Trailblazers with thirty-seven, while the Orlando Magic had the least with just fifteen.

Table 1 presents the final percent ASID metric for each NBA team. The Golden State Warriors are the team with the highest spacing impact by our metric, with a percent ASID of 21.74%. In comparison, the average percent ASID for the league is -0.63% and the Memphis Grizzlies have the lowest percent ASID with -15.97%.

Figure 2 provides a visual comparison of percent ASID throughout the league. Figure 3 plots the percent ASID versus the number of three-point attempts per 100 possessions for the 2015-16 season.[6]

[6]Three-point data from basketball-reference.com.

Table 1: Percent ASID for each NBA team, split by conference during the 2015-16 season

Eastern Conference		Western Conference	
Team	ASID (%)	Team	ASID (%)
ATL	-0.41	DAL	-8.09
BKN	-8.59	DEN	1.73
BOS	-1.28	GSW	21.74
CHA	0.41	HOU	-2.05
CHI	-5.25	LAC	1.55
CLE	1.53	LAL	0.54
DET	-1.86	MEM	-15.97
IND	-2.06	MIN	-6.63
MIA	-3.27	NOP	-5.54
MIL	-9.29	OKC	15.60
NYK	9.27	PHX	2.52
ORL	-5.59	POR	2.17
PHI	10.85	SAC	8.33
TOR	-7.78	SAS	-10.79
WAS	3.86	UTA	-4.46

Figure 2: Percent ASID for each team in the NBA

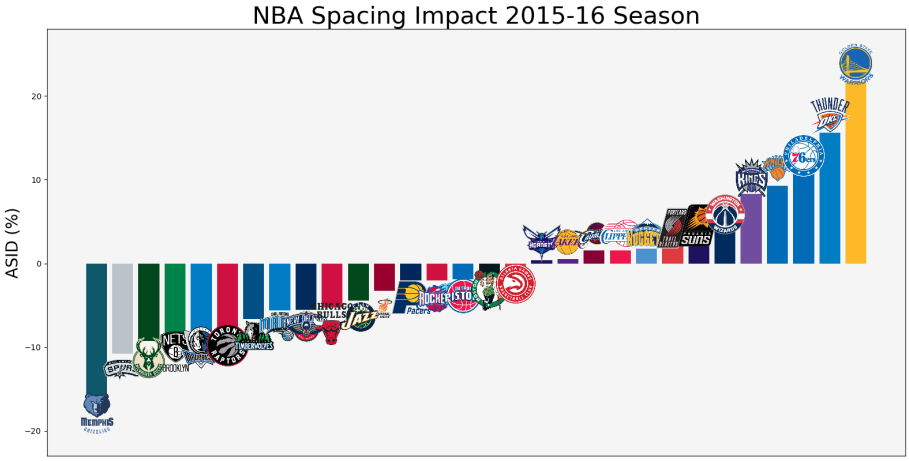
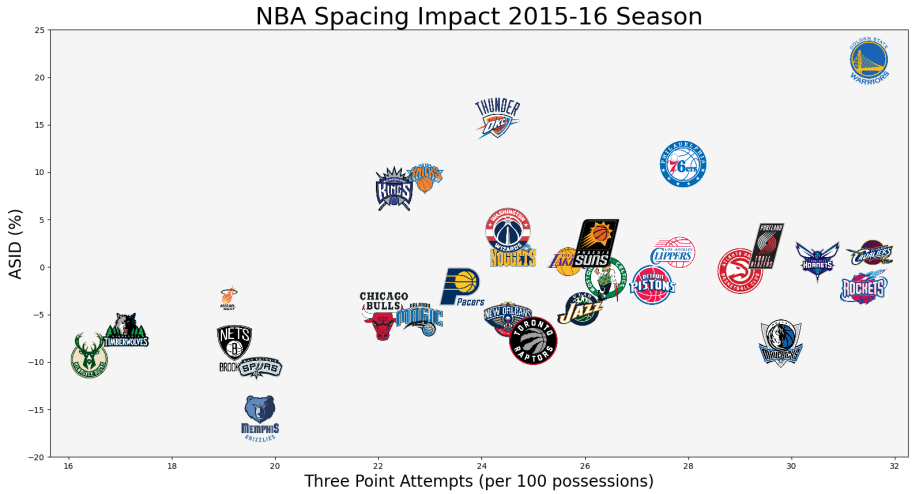


Figure 3: Percent ASID versus three-point attempts per 100 possessions



5 Discussion

The ASID metric allows us gain valuable insights into the impact certain teams have on opponent's defensive spacing. While this impact may be observable by viewing game footage, the ASID metric allows us to precisely quantify it.

The obvious standout application of the metric is the case of the Golden State Warriors. The Warriors are the team with the highest ASID by a large margin. This is likely due to their overwhelming three-point shooting during the 2015-16 season,[7] which forced defenses to guard closer on the perimeter.

The metric also provides less obvious insights, such as with the Toronto Raptors. Despite having the fourth highest three-point percentage in the league during the 2015-16 season, their percent ASID was -7.78%. From this we can infer that despite having success with three-pointers, they were not able to effectively stretch the defense. This could be due to their offensive strategy, or simply that they did not shoot enough three-pointers.

Analysis using the specific results obtained in this paper may be limited due to the dataset. Factors such as the uneven number of games used for each team, as well as the variable number of usable shots within each game, likely had an effect on the final results. However, the overall efficacy of the metric is promising. Future work to further evaluate the metric has merit as the metric seems to quantify spacing impact well.

[7] The 2015-16 Warriors made 1077 three-pointers during the regular season, a record at the time.

6 Conclusion

In this paper I propose a novel metric called Average Spacing Impact on the Defense (ASID). The metric allows for quantification of a team's impact on the spacing of the defense. Effectively this metric quantifies how well a team is "spacing the defense." This provides valuable feedback on the execution of a team's offensive scheme and can be used for further analysis of team performance. While the results obtained in this paper are limited due to lack of data, this metric can be easily extended to an entire NBA season, or could be used to track a team's spacing throughout a season, assuming tracking data is available.

7 References

Bourbousson J, Sève C, McGarry T (2010) Space–time coordination dynamics in basketball: Part 2. The interaction between the two teams, Journal of Sports Sciences, 28:3, 349-358, DOI: 10.1080/02640410903503640

Franks A, Miller A, Bornn L, Goldsberry K (2015) Characterizing the spatial structure of defensive skill in professional basketball, The Annals of Applied Statistics, 9:1, 94–121, DOI: 10.1214/14-AOAS799

Frencken W, Lemmink K (2008) Science and Football VI, Routledge, Chapter 27, Team kinematics of small-sided soccer games: A systematic approach, pp.161-166

Frencken W, Lemmink K, Delleman N, Visscher C (2011) Oscillations of centroid position and surface area of soccer teams in small-sided games, European Journal of Sport Science, 11:4, 215-223, DOI: 10.1080/17461391.2010.499967

Metulini R, Manisera M, Zuccolotto P (2017) Sensor Analytics in Basketball. Proceedings of the 6th International Conference on Mathematics in Sport. ISBN 978-88-6938-058-7