Beyond the Numbers: A Visual Display of Advanced NBA Defensive Analytics

Stephen Pelkofer
Georgia Insitute of Technology

Richard Reasons
Georgia Institute of Technology

Matt Luppi
Georgia Institute of Technology

Hyun Jin lee
Georgia Insitute of Technology

Tim Kim
Georgia Insitute of Technology

1 THE PROPOSAL

Sports analytics have provided ways for fans of all sports to become more engaged with their favorite games. Unfortunately for fans, analytics have mainly focused on the offensive aspects of the game. Our team's aim is to provide an interactive web page that will allow fans of the National Basketball Association (NBA) to explore, visualize, and assess the *defensive* performance of the league's teams and players.

2 TODAY'S APPROACH

2.1 Mundane Presentation of Data

The NBA has a very active, analytical following that pores over team performance, player statistics, and salary cap implications. Websites such as NBA Advanced Stats [1], Cleaning the Glass [2], and FiveThirtyEight [3] provide fans with some raw statistics, but predominantly focus on player offensive performance. While there have been some attempts to enhance defensive stats, such as efficiency of rebounding [4] and forcing weak side dribble [5], it is limited in its applications. Typically, data is presented through tabular formats or uninteresting visualizations, offering fans plenty of data, but little insight.

2.2 A Focus on Offensive Metrics

Shot charts and box scores have been available to fans for decades, but the NBA now exposes granular play-by-play data with spatial player tracking data. Despite its recent introduction, this data has already been utilized in exciting ways for new offensive metrics: e.g. assessing shooting effectiveness from certain areas of the court [6] and identifying player's habits when it came to how they positionally shot the basketball [7]. Due to the dearth of quantifiable statistics, defense has been historically difficult to measure.

3 OUR TAKE

Therefore, there is clearly a need for analyses that focuses on quantifying and visually capturing basketball defensive efficacy. The NBA's granular play-by-play tracking data introduce the possibility for exciting defensive metrics. Some research has already been conducted using this data for new defensive metrics, proposing layouts for innovative defensive shot charts [8], interrelated clusters of current metrics [9], and insight into how the presence of other players is enough to deter shooters [10].

These metrics have yet to appear in any leading NBA coverage sites. The statistical methods are available, but there is no presentation for fans of the game. We believe we can follow the framework and guidance presented by Sleeper[11], Yaucite [12], and Tufte[13] to create an intuitive dashboard presentation. This will offer insight to new defensive metrics through interactive graphics. After all, humans process information quicker when data is presented through visually compelling schema.

4 WHO CARES?

The research we were able to uncover shows that there is an appetite for basketball analysis. Our team is proof of this belief, brought together by a shared love of the sport, but we know we are not alone. The reddit NBA

[14], a hub of hardcore NBA fandom, has over 3 million members. Conversations around analysis by the previously cited sites such as FiveThirtyEight[15] garner thousands of upvotes and hundreds of comments. Providing a way for these fans to analyze the neglected half of the sport would spur new conversations and injects new talking points into fan discussions.

Additionally, this data could even be useful to NBA teams or basketball teams at any level looking for an edge to improve their defensive schematic.

5 IMPACT AND MEASUREMENT

There is a big opportunity to create a robust defensive analytics dashboard for the NBA community. Fans, blogs, website, and possibly even teams will be able to use our analysis to gain insight into defensive performance. The main methods for measurement of impact will be how many views are website gets and how many times are website is mentioned in forums such as twitter, reddit, and other websites.

6 RISKS AND PAYOFFS

The primary risks of the project have to do with potential scope creep. The NBA provides some defensive data via an API, but we will likely have to scrape this data ourselves to create a dataset fit for robust analysis. Even then, a considerable amount of work will be required to translate the numbers via a method published in a journal to a visual display on a web page. However, the payoffs from this project would outweigh those risks, filling a blind spot in the NBA community for both fans and teams.

7 COSTS

We anticipate the monetary costs of the project to be minimal, totaling less than $200 dollars for the life of the project. The primary output, an interactive web visualization, will only cost a few dollars a day in web hosting. The only other expense would be the virtual machine instances, if necessary, to scrape our web sources and aggregate the raw data. Renting these instances for the few days necessary to scrape our dataset will cost less than ten dollars based on pricing calculators for Microsoft Azure [16] and Google Cloud Platform [17].

8 TIMELINE AND PLAN OF ACTIVITIES

All team members have contributed a similar amount of effort to this point.

- Feb 28: Project proposal (all)
- Mar 27: Progress report (all)
  - Feb 28 - Mar 7: Scrape data (S.P.)
  - Feb 28 - Mar 14: Wireframe visuals and design dashboard (T.K., M.L., H.L., R.R.)
  - Feb 27 - Mar 14: Build frontend of website with R Shiny (S.P., R.R.)
  - Mar 14 - Mar 27: Incorporate D3-based visuals (T.K., M.L, H.L.)
- Apr 17: Final Report (all)
  - Mar 27 - 4 Apr: Build parallel paths to build user-facing site (all)
  - 4 Apr - Apr 17: Add remaining visuals (all)
  - 4 Apr - Apr 17: Rigorously test functionality (all)

9 PROGRESS CHECKS

The project's success will be evaluated via a final report, poster, and presentation. Additionally, we will share a work-in-progress draft of our final report partway through the project, including any setbacks we have encountered or changes in scope that we have had to undertake. We plan to conduct the project in an iterative nature, building subsets of the final project piece-by-piece as we go along. By doing so, we can measure success as the amount of our final project components that have been completed, reviewed, and deployed to our testing environment.

10 CONCLUSION

In Sprawlball, Goldsberry references that the recent advancements of computing and collecting spatial data and this generation of sports fan's obsession with quantification and efficiency has led to a widespread adoption of analytical analysis in the NBA [18]. Advanced statistics are a big talking point in most talk-shows, podcasts, and even the in-game commentators. There is a large audience that is interested in quantifying the game, and yet not much has been accomplished for quantifying defense. Our project will provide fans across the world an additional resource to leverage when analyzing and assessing their favorite sport.

REFERENCES

[1] Nba advanced stats. http://stats.nba.com.

[2] Cleaning the glass. http://www.cleaningtheglass.com.

[3] Fivethirtyeight. http://www.fivethirtyeight.com.

[4] James T. Bartholomew and David A. Collier. The benefits of forcing offensive basketball players to their weak side. *Journal of Multidisciplinary Research*, 4:19–27, 2012.

[5] et al Simovic, Slobodan. Analysis of influence of basketball game-related statistics on final result based on differences at the 2017 fiba asia cup. *Asian Journal of Sports Medicine*, 10.1, 2019.

[6] Bradley A Sliz. An investigation of three-point shooting through an analysis of nba player tracking data. *arXiv preprint arXiv*, 1703.07030, 2017.

[7] et al Franks, Alexander. Characterizing the spatial structure of defensive skill in professional basketball. *The Annals of Applied Statistics*, 9.1:94–121, 2015.

[8] et al Franks, Alexander. Counterpoints : Advanced defensive metrics for nba." mit sloan sports analytics conferencel. *MIT Sloan Sports Analytics Conference*, 2015.

[9] et al Franks, Alexander. Meta-analytics: tools for understanding the statistical properties of sports metrics. *Journal of Quantitative Analysis in Sports*, 12.4:151–165, 2016.

[10] Kirk Goldsberry and Eric Weiss. The dwight effect: A new ensemble of interior defense analytics for the nba. *Sports Aptitude, LLC*, pages 1–11, 2013.

[11] Ryan Sleeper. Practical tableau: 100 tips, tutorials, and strategies from a tableau zen master. *O'Reilly Media, Inc*, 2018.

[12] Nathan Yau. Visualize this: The flowingdata guide to design, visualization, and statistics. *Wiley*, 2019.

[13] Edward Tufte. *Edward. The Visual Display of Quantitative Information.*

[14] Reddit. https://www.reddit.com/r/nba/comments/ce17l7/fivethirtyeights_201920_nba_projections/.

[15] Fivethirtyeight's 2019-20 nba projections. https://projects.fivethirtyeight.com/2020-nba-predictions/.

[16] Pricing calculator: Configure and estimate the costs for azure products. https://azure.microsoft.com/en-us/pricing/calculator/.

[17] Google cloud platform pricing calculator. https://cloud.google.com/products/calculator.

[18] Kirk Goldsberry. Sprawlball. sprawlball: A visual tour of the new era of the nba. houghton mifflin harcourt. 2019.