Recognizing and Analyzing
Ball Screen Defense in the NBA

Avery McIntyre[1], Joel Brooks[2], John Guttag[2], and Jenna Wiens[1]

[1]Computer Science & Engineering
University of Michigan,
Ann Arbor, MI 48104
mcintyav, wiensj @umich.edu

[2]Computer Science and Artificial Intelligence Lab
Massachusetts Institute of Technology
Cambridge, MA, 02139
guttag, brooksjd @mit.edu

Abstract

As the NBA's go-to offensive play, determining how to defend the ball screen is among the most critical decisions faced by NBA coaching staffs. In this paper, we present the construction and application of a tool for automatically recognizing common defensive counters to ball screens. Using SportVU player tracking data and supervised machine learning techniques, we learn a classifier that labels ball screens according to how they were defended. Applied to data from four NBA seasons, our classifier identified 270,823 screens in total. These labeled data enable novel analyses of defensive strategies. We present observations and trends at both the team and player levels. Our work is a step towards the construction of a coaching assistance tool for analyzing one of the game's most important actions.

1   Introduction

With half of Stephen Curry's 25 three-pointers during the 2015 NBA Finals involving pick and rolls, finding an effective counter to the ball screen clearly remains a vexing matter for defenses. The momentary havoc caused by the screen requires coordination to impede the ball handler, deny an opportunity for the screener, and avoid an untenable mismatch. Mismanagement of this maneuver often results in an offensive advantage and a score.

In this paper, we present a step towards the construction of a tool designed to help coaching staffs analyze the way ball screens are defended. It can be used to automatically identify the way ball screens are defended, and measure correlations between defensive strategies and possession level outcomes.

Utilizing SportVU player tracking data [1] from multiple seasons in the NBA and supervised machine learning techniques, we learn to accurately classify pick and roll defensive schemes as "over," "under," "trap," or "switch." We then apply our learned classifier to multiple seasons worth of data, consisting of a total of 270,823 ball screens, generating labels for how each ball screen was defended. Combined with data about outcomes (e.g., points per possession), these labeled data enable novel analyses at the team and player level. We examine the distributions of schemes for teams across the league, and then turn our analysis to individual players examining the effect of individual schemes on offensive and defensive performance.

1

2 Data

We consider player trajectory data collected by the STATS SportVU system. This dataset contains all (x, y) positions of every player on the court and the (x, y, z) coordinates of the ball at 25 frames per second. These trajectory data are augmented with play-by-play data including player information, box scores, game clock, shot clock etc. We consider data from four separate seasons summarized in Table 1. We have fewer data points for earlier seasons, since the system was not installed across the League until 2013. In our analyses, we focus on data from 2013-2014, since that is the season for which we have the most complete data. When examining overall trends, or trends over time we consider all four seasons.

Table 1: The amount of available data for analysis varies across seasons.

Season	Total Games
2011-2012	254
2012-2013	611
2013-2014	1315
2014-2015	716

3 Learning to Classify Defensive Schemes

Building upon earlier work on recognizing the occurrence of an on-ball screen [2], we develop a system that takes as input unlabeled player trajectory data from SportVU and outputs the time of all ball screens that occur during each game and how the screen was defended. Figure 1 illustrates the three stages to the overall system. In this paper, we focus mainly on the third stage in which we classify defensive schemes. For details on the first two stages, please refer to [2].

E.g., [clock=11:55, player_id=230 , x=34.3, y=25.4]
[clock=11:55, player_id=245 , x=30.3, y=20.4]
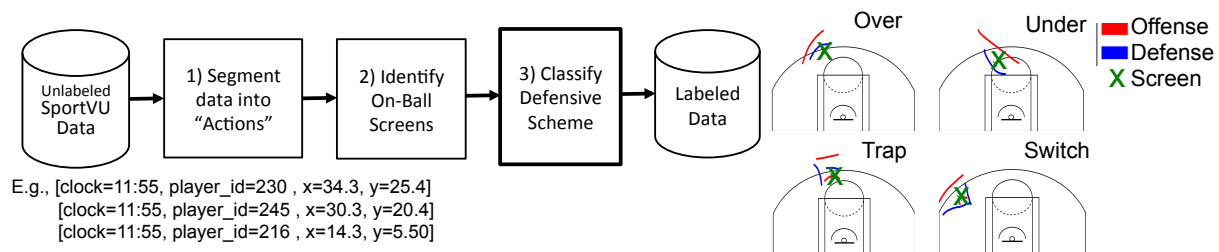[clock=11:55, player_id=216 , x=14.3, y=5.50]

Figure 1: Our system takes in SportVU position data and identifies ball screens and labels them with the defensive scheme used and the offensive outcome in terms of points per possession.

3.1 Definitions

For each instance of a ball screen, we identify the following four players: the *ball handler*, the *on-ball defender*, the *screener*, and the *screener defender*. We group the defensive schemes into four broad categories based on the trajectories of these players:

Over: On-ball defender stays between the ball handler and screener i.e., goes "over" the screen
Under: On-ball defender does not stay between the ball handler and screener i.e., goes "under"
Switch: On-ball defender and screener defender switch their original matchups
Trap: On-ball defender and screener defender double the ball handler i.e., "trap" the ball-handler
Within each of these broad categories, there are subtypes of interest e.g., "hedge" (in which the screener defender jumps out on the ball handler to impede his progress), or "show" (in which the screener defender slides out just far enough to present an obstacle to the ball handler). We do not present results at this level of detail here.

3.2 Learning the Classifier

In the third stage of the pipeline in Figure 1, we classify each screen according to how the defense behaved. Rather than hand-coding a set of rules, we used supervised learning to automatically build models that can be used to classify defensive activity as one of the four categories listed above. To build our training set of data, we watched film from six games from the 2012-2013 regular season. In total we hand labeled a set of 340 attempts to defend a ball screen. Each attempt was labeled based on its most defining characteristic. For example, we would label an instance where the on-ball defender goes over and then a trap occurs as a trap, since the occurrence of a trap defines the screen defense. In total our training set consisted of 199 instances of over, 56 instances of under, 57 instances of switch, and 28 instances of trap.

For each example, we extracted features based on the pairwise distances between the four player trajectories. We extracted each segment automatically using our ball screen detector. The examples varied in length, i.e. we have some examples that contain more samples than others. The largest variance is between the time the play begins and when the ball-handler reaches the screener. To align the examples in time, we define the "screen moment" as the when the distance between the screen and the on-ball defender is minimum. We align all examples based on this moment and consider ten frames prior to this moment. We consider the player trajectories from this moment up until the moment a shot, pass, turnover, or stoppage of play occurs. For each of the n examples in our dataset, we represent each of the pairwise distances by a variable length vector $z_i \in \mathbb{R}^{d_i}$ where $i = 1...n$. From these variable length vectors, we extract summary statistics describing the signal (e.g., the mean, the maximum, the minimum, etc.). Concatenated together, these features result in a fixed length feature vector $\mathbf{x} \in \mathbb{R}^d$ for each example.

To train and validate our model, we split our hand labeled data 70/30 into a training set and validation set. Since the schemes are not uniformly represented in our data, we consider stratified splits that maintain the overall label/class distribution. Given the training data, we define a min/max range, and scale the validation set features accordingly to ensure that all features have the same scale. We then learn a linear multiclass classifier using multinomial logistic regression [3]. We select the hyperparameters using 5-fold cross validation on the training set. In addition, we account for the non-uniform distribution across classes by adjusting class weights to be inversely proportional to class frequencies [4]. The learned classifier results in a mapping from a feature vector $\mathbf{x}$ representing a ball screen to a class (i.e., label) corresponding to one of the four defensive schemes.

3.3 Classifier Performance

We evaluate the classification performance of our learned classifier on the labeled test data. Given an unlabeled example $\mathbf{x}$, our classifier produces a probability estimate for each class j:

$$P(y = j|\mathbf{x}) = \frac{e^{\mathbf{x}^T \mathbf{w}_j}}{\sum_{k=1}^4 e^{\mathbf{x}^T \mathbf{w}_k}} \quad \text{and} \quad \hat{y} = \underset{j}{\mathrm{argmax}} P(y = j|\mathbf{x})$$

where $\mathbf{w}_j$ are the learned regression coefficients for classifier j. We assign each example a label, $\hat{y}$, according to the class that results in the maximum probability. We repeat this process of splitting the

data, training a classifier, and testing the classifier on 100 different splits. For each split, we measure performance in terms of the precision and recall with respect to each class, in addition to overall accuracy. We achieve an overall average accuracy of 0.69 ($\pm$0.03). Table 2 lists the performance within each class and Table 3 gives the confusion matrix averaged over the 100 different training/test splits. Among the four different schemes, the classifier has the greatest difficulty classifying traps. Trap is also the least represented in the training set.

Table 2: Average performance ($\pm$std. dev.) on validation sets within each class.

Metric	Over	Under	Trap	Switch
Recall	0.83($\pm$0.05)	0.52($\pm$0.12)	0.19($\pm$0.13)	0.62($\pm$0.11)
Precision (with $h = 0$)	0.75($\pm$0.03)	0.58($\pm$0.09)	0.40($\pm$0.25)	0.63($\pm$0.09)
Precision (with $h = 0.6$)	0.78($\pm$0.05)	0.65($\pm$0.12)	0.46($\pm$0.36)	0.69($\pm$0.11)

Table 3: Confusion matrix averaged over 100 random cross-validation splits

		True Class Label			
		over	under	trap	switch
Predicted	over	**49.64**	5.63	6.50	4.56
Class	under	4.24	**8.83**	0.19	2.02
Label	trap	2.16	0.24	**1.72**	0.25
	switch	3.96	2.30	0.59	**11.17**

In order to increase performance, and specifically our ability to precisely classify traps, we considered a second classification scheme based on a threshold h such that:

$$\hat{y} = \begin{cases} \underset{j}{\operatorname{argmax}} P(y = j | \mathbf{x}) & \text{if} \quad \underset{j}{\max} P(y = j | \mathbf{x}) > h \\ \text{unclear} & \text{otherwise} \end{cases}$$

By thresholding how confident the classifier must be in its prediction, we can increase overall precision. This results in a reduction in recall since an example may not be labeled as belonging to any class, however in some analyses we care more about precision than recall so we are willing to accept this tradeoff.

Next, we applied our classifiers to all 270,823 ball screens previously extracted by stages 1 & 2. This results in 146,314 examples of "over", 69,721 of "under", 37,336 of "switch", and 17,451 of "trap". When we set $h = 0.6$ to increase precision, we identified 33,421 examples of "over", 9,252 of "under", 8,394 of "switch", 383 of "trap", and 219,372 examples are labeled unclear. We analyze these labeled data and report our observations in the next section.

4   Observations and Trends in Defensive Schemes

Applied to four seasons worth of data, our ball screen classifier identified 270,823 ball screens. In addition to knowing the outcome for the possession (e.g., number of points scored), we also know who was involved. These data allow for an unprecedented analysis into the effectiveness of defensive

strategies. Here, we begin to analyze these data and identify trends in defensive strategies at both the team and player levels.
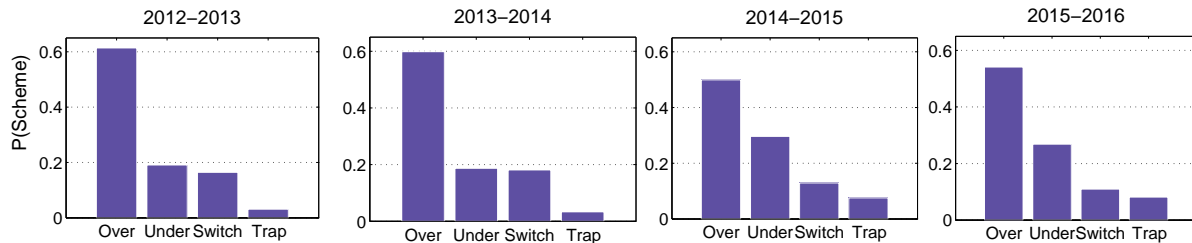
4.1 Team-Level Analysis

Figure 2: Across seasons the observed distribution stays relatively constant, however we do notice a slight (though probably not significant) increase in the number of traps.

As a preliminary analysis we looked at the general defensive tendencies of teams by examining the probability distribution for each team across the four schemes, see Figure 2. Because our classifier does not achieve 100% accuracy, we consider relative differences across teams and seasons. From season to season, the average distribution of schemes is reasonably consistent. The only noticeable difference is an increase in the frequency of "under." We compared teams to each other by calculating the difference between the distributions using symmetric version of the Kullback-Leibler divergence. These pairwise differences are given in Figure 3(c). We also compared teams to the average distribution for all teams in order to identify outliers. Figures 3(a) and 3(b) show how much teams' defensive scheme distributions varied from average during the 2013-2014 season. These plots also illustrate how a team's strategy for defending screens relates to winning percentage over the course of a season. Larger, lighter dots imply a higher regular season win percentage. Some teams are clear outliers. The 2013-2014 Miami Heat trapped nearly twice as frequently (13.2% of opponent screens) as the average team that season (just 7.5%). This is probably not unrelated to their forcing of the second most turnovers of any team that season. The 2013-2014 Bulls used "over" at a rate almost 7% greater than the average team that year. We believe this was an important aspect of the team's overall defensive scheme of forcing ball handlers towards the sidelines.

4.2 Player-Level Analysis

In this section we examine how personnel groups, both offensive and defensive relate to the way screens are defended and how these decisions result in possession level outcomes (i.e., points per possession).
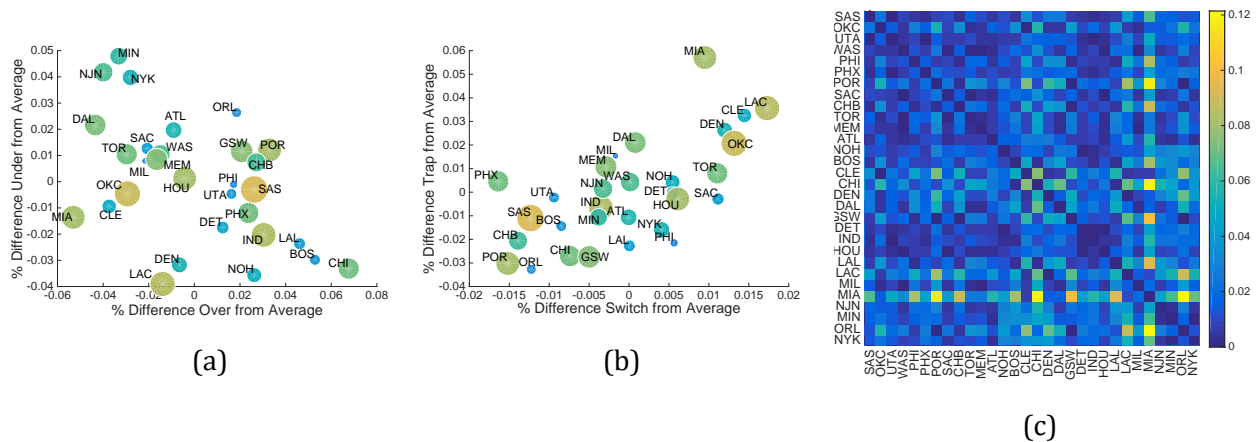
Figure 3: In 2013-2014 (a) The Bulls went "over" almost 7% more often than the average team . (b) While the average team "trapped" just 7.5% of ball screens, the Heat "trapped" more than 13% of the time. (c) The difference between defensive strategies of pairs of teams (lighter=greater difference). By this metric, the Clippers and Heat were among the most unique defensive teams.

4.2.1 Offense

It is well known that how teams choose to defend a ball screen is strongly related to the offensive skill set (particularly the shooting ability) of the ball handler. To support this claim with data, we applied our classifier ($h = 0.6$) to all data collected from 2012-2015. For this analysis, we set the classifier threshold (h) higher, so that our precisions among each defensive scheme was at its highest. This resulted in approximately $51,000$ ball screens labeled by their defensive scheme. Only about 350 of those instances were labeled as traps, so we did not include traps as part of this analysis.

We considered ball handlers with more than 300 labeled instances of ball screens. For each of these 53 players, we compute the distribution of how teams chose to defend ball screens when that player was the ball handler. For each defensive strategy, we also compute the average points per possession across all ball screens in which that player was the ball handler. For example, we calculate that, on average, the offense scores 0.83 points every time the defense switches on LeBron James.

Figures 5-7, show the effectiveness of over, under, and switch against these 53 players. The x-axis is the percentage of time the strategy is used, and the y-axis the average points per possession the offense scores when that strategy is used. Players near the upper-right quadrant of a plot are players that are defended
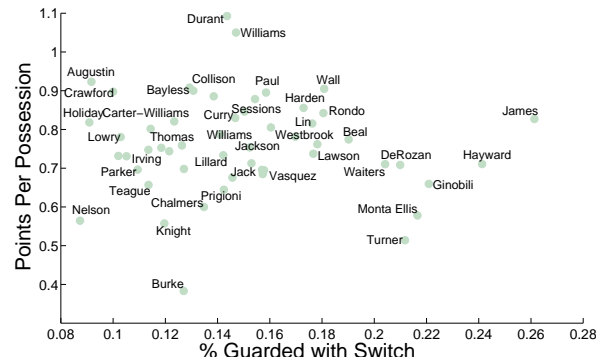

Figure 4: Avg. Points/Possession when guarded with Switch versus relative frequency of being guarded with Switch. While faced with the switch more often than any other player, James remains among the most effective at creating points. Durant punishes the defense to an even greater extent, but encounters the switch less often.

often by a particular scheme, while still managing to score. Players in the lower-right quadrants also see that particular scheme often, but are not as effective at defeating it.
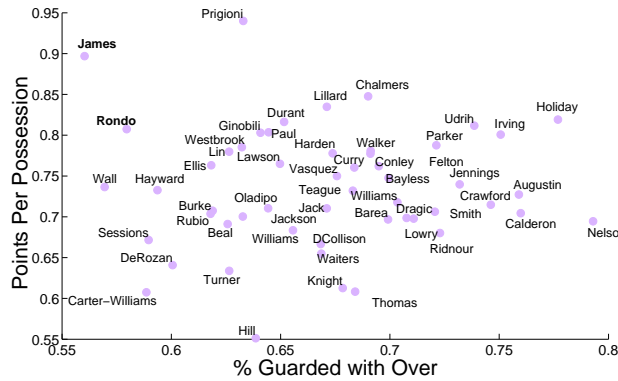
Figure 5: Avg. Points/Possession when guarded with Over versus relative frequency of being guarded with Over. Prigioni, Chalmers, and Holiday are among the players that excel in creating points even when the defense frequently goes over their screens.
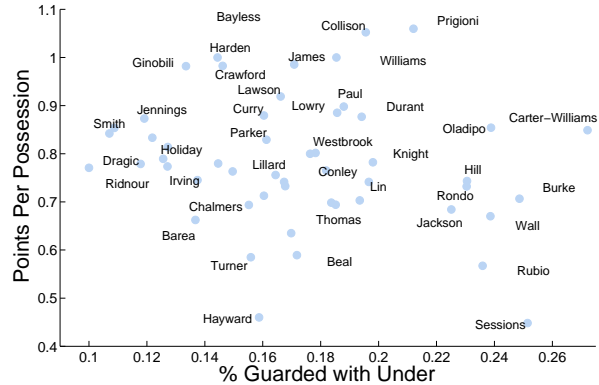

Figure 6: Avg. Points/Possession when guarded with Under versus relative frequency of being guarded with Under. In general, players who consistently make the three-point shot tend to dominate when the defense goes under.

Against over, we see that players like James, Prigioni, Chalmers, Lilliard, and Holiday all result in a relatively high number of points per possession. In spite of this success, teams still choose to defend Irving and Holiday by going "over" more than 75% of the time.

For under, we again see Prigioni and James near the top in terms of points per possession. When Michael Carter-Williams is the ball handler, the defense goes under the screen more often than against others. Still, Michael Carter-Williams is relatively successful against this defensive tactic, while Hayward and Sessions fair poorly when facing the under strategy.

Against switches, Durant and Mo Williams lead to more points per possession than others. Of all the 53 players we investigated, LeBron James faced the switch most often. However, James is slightly more effective when the defense switches (vs. over/under) suggesting that switch as the strategy of choice when defending against LeBron may be suboptimal. The switch is, however, more effective when defending against Chalmers and Prigioni relative to going over.

4.2.2 Defense

In addition to analyzing ball screen defense from the perspective of the offense, we can flip the lens and examine how defensive success varies across different defensive players.

First, we consider the most effective defense (i.e, the pair of defensive players that yield the fewest points per possession) for each type of scheme using data from the 2013-2014 season. The top ranked defensive pairs are shown broken down by defensive scheme in Table 4.

Note that some particularly effective pairs/players do not appear in these tables since we did not have enough samples. We consider only those pairs for which we have a large enough number of samples to make meaningful observations.

Table 4: 2013-2014: Most effective teams at defending the ball screen within each defensive scheme.

Team	Defensive Player 1	Defensive Player 2	Total (Fraction)	Avg. Pts/Possess.
		Over		
Thunder	Derek Fisher	Steven Adams	14 (0.58)	0
Thunder	Reggie Jackson	Steven Adams	20(0.69)	0
Thunder	Thabo Sefolosha	Serge Ibaka	15 (0.45)	0.13
Knicks	Carmelo Anthony	Raymond Felton	15 (0.56)	0.25
Hawks	Jeff Teague	Pero Antic	19 (0.56)	0.26
Rockets	Dwight Howard	Jeremy Lin	40 (0.73)	0.27
76ers	Spencer Hawes	James Anderson	14 (0.74)	0.28
Knicks	Tyson Chandler	Iman Shumpert	14 (0.56)	0.28
Heat	Chris Bosh	Norris Cole	18 (0.50)	0.33
Suns	Eric Bledsoe	Channing Frye	18 (0.82)	0.33
		Under		
Nets	Shaun Livingston	Andray Blatche	20 (0.47)	0.20
Raptors	Kyle Lowry	Jonas Valanciunas	13 (0.23)	0.31
Rockets	Dwight Howard	James Harden	12 (0.32)	0.31
Magic	Tobias Harris	Jameer Nelson	13 (0.35)	0.31
Heat	Norris Cole	Chris Anderson	14 (0.28)	0.36
76ers	Spencer Hawes	M. Carter-Williams	13 (0.32)	0.38
Trail Blazers	Nicolas Batum	Damian Lillard	13 (0.38)	0.46
Knicks	Raymond Felton	Tyson Chandler	17 (0.33)	0.47
Bulls	D.J. Augustin	Taj Gibson	14 (0.26)	0.50
Bulls	Carlos Boozer	Kirk Hinrich	12 (0.23)	0.50
		Switch		
Cavaliers	Jarrett Jack	Anderson Varejao	13 (0.39)	0.38
Clippers	Chris Paul	Blake Griffin	19 (0.25)	0.42
Raptors	Kyle Lowry	Amir Johnson	16 (0.22)	0.50
Wizards	Marcin Gortat	Bradley Beal	11 (0.39)	0.64
Heat	LeBron James	Mario Chalmers	11 (0.15)	0.73
Wizards	Marcin Gortat	John Wall	15 (0.15)	0.73
Pacers	George Hill	David West	13 (0.13)	0.77
Grizzlies	Mike Conley	Zach Randolph	15 (0.21)	0.8
Heat	Chris Bosh	Mario Chalmers	16 (0.25)	0.81
Heat	Chris Bosh	Dwayne Wade	11 (0.35)	0.91

Interestingly, pairs that defend well with one scheme do not appear dominant across all schemes (or simply do not run the same defense). From this, we can note the difficulty that arises when attempting to determine the overall defensive effectiveness of a single player. Our analysis suggests that the defenders vary considerably in their ability to thwart different ball screen actions by the offense.

We can determine individual defensive effectiveness by examining all ball screens in which a given player was among the pair of defensive players. When we consider each individual player's average performance across all pairs in this way, some clear winners come out on top. In particular, Steven Adams is a standout defender; regardless of who he is paired with, Adams and his teammate are particularly effective at blocking the defense from scoring off of a pick and roll. Similarly, Varejao and Howard are standouts within other categories.

Figure 7 compares the average points per possession allowed by different duos to the average allowed by each individual when averaged across all other player combinations.

Chris Paul and Blake Griffin, while effective at the switch (see TABLE 4), are one of the worst pairs when it comes to "over", averaging close to 1.2 points allowed. Based on each individual's contribution, Griffin appears to be on the hook for this one. Similarly, Serge Ibaka and Reggie Jackson perform relatively poorly at the switch, but appear equally responsible.
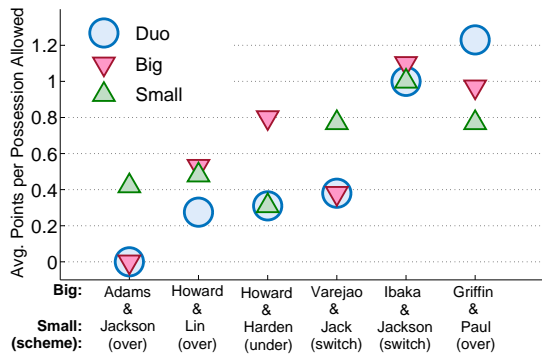
Figure 7: Points per possession allowed for various duos and individuals. While Howard and Lin are particular synergistic in defending the pick and roll, Blake Griffin and Chris Paul perform worse together than they do in other pairs.
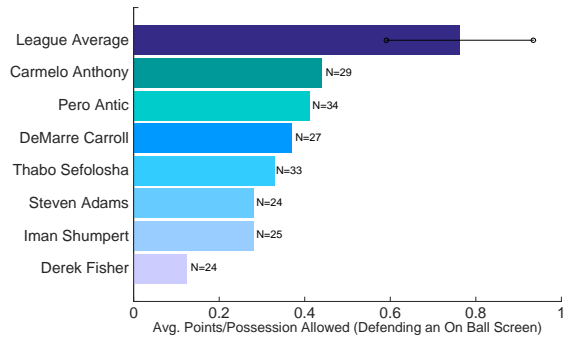

Figure 8: Avg. Points/Possession allowed by players who were significantly better than the League average. Notably, Shumpert, Adams, Sefalosha, and Carroll are all well regarded as solid defenders.

When averaged across all schemes the players who come out on top in terms of most effective defense (when defending the pick & roll) are shown in Figure 8. Compared to the league average these players are more effective by two or more standard deviations.

In the most recent season 2014-2015, we notice slightly different trends. In terms of "over", Westbrook and Ibaka perform best compared to all other pairs, with Joakim and Rose a close second. Interestingly, Rose and Pau Gasol are one of the worst teams at running "over". Again, based on individual player contributions, Ibaka appears to be carrying the Westbrook/Ibaka duo. Joakim and Rose appear to be particularly synergistic; paired with others, neither do very well, but together they rank among the top defenders.

In terms of under Paul and Griffin perform similarly to 2014, with one of the worst performances relative to the other pairs (again Griffin appears to be on the hook). Pau Gasol and Derrick Rose are also ineffective
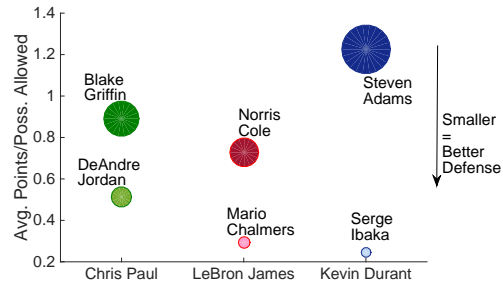

Figure 9: Avg. Points/Possession allowed for various players and the teammates they most frequently defend with. Paul, James and Durant all have a teammate with whom they defend ball screens especially well.

when running under. Overall, Paul and Griffin, Gasol and Rose, and LaMarcus Aldridge and Damian Lillard are some of the most frequent defensive duos with the worst performance. In 2015, Serge Ibaka was a highly effective defensive player, allowing on average only 0.25 points per possession when defending the pick and roll.

Across all seasons some noteworthy combinations appear, Chris Paul and DeAndre Jordan, Durant and Ibaka, and James and Chalmers. Figure 9 illustrates the effectiveness of these combinations in addition to others that appear frequently in the data but result in worse performance.

2016 Research Papers Competition
Presented by:

5 Conclusion and Potential Applications

The classification process that we present for identifying defensive schemes helps discretize the game in a way that offers value beyond summary statistics. Our process allows for greatly improved insight into the pick and roll, for which we see at least two potential applications:

1. **Player-Level Reporting**
   With the ability to automatically identify pick and roll actions and responses, our process moves towards a system for individualized, player-level reporting tools. Such tools could be used to quickly and automatically assess a player's performance during a game on actions of particular interest and serve as a guide during player development work.

2. **Strategic Planning**
   When assessing team performance over a series of games, our process could be applied in order to assist in strategic planning. Teams could immediately identify areas of weak defensive performance and prioritize this information in preparing for future games.

In continued analysis, we also see potential to assess evolutions in the defensive styles of teams across the League, as well as examine how the defensive abilities of an individual player affect overall outcomes for the team in pick and roll situations.

References

[1] http://www.stats.com/sportvu/sportvu.asp. STATS sportVU, 2015.

[2] Armand McQueen, Jenna Wiens, and John Guttag. Automatically recognizing on-ball screens. In *2014 MIT Sloan Sports Analytics Conference*, 2014.

[3] Kevin P Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

[4] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, et al. Scikit-learn: Machine learning in python. *The Journal of Machine Learning Research*, 12:2825–2830, 2011.