Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v10n3p712

Role revolution: towards a new meaning of positions in basketball

By Bianchi, Facchinetti, Zuccolotto

Published: 15 November 2017

Electronic Journal of Applied Statistical Analysis
Vol. 10, Issue 03, November 2017, 712-734
DOI: 10.1285/i20705948v10n3p712

Role revolution: towards a new meaning of positions in basketball

Federico Bianchi[a], Tullio Facchinetti*[a], and Paola Zuccolotto[b]

[a] *University of Pavia, Dept. of Industrial, Computer and Biomedical Engineering, via Ferrata, 1, 27100 Pavia, Italy*
[b] *University of Brescia, c.da S. Chiara 50, 25122 Brescia, Italy*

Published: 15 November 2017

Most team sports are characterized by positions to which the players in a team are assigned. The goal of this classification is to attribute specific responsibilities during a game. Moreover, the same classification drives the buying and selling of players according to team managers and coaches strategies. The existing positions - often defined a long time ago - tend to reflect traditional points of view about the game and sometimes they are no longer well-suited to the new concepts arisen with the evolution of the way of playing.

This paper focuses on basketball and aims at describing new roles of players during the game, by means of the analysis of players' performance statistics with data mining and machine learning tools. In detail, self-organizing maps and fuzzy clustering procedures are adopted in tandem to define groups of players with similar way of playing. The results show that, when considering the modern basketball players' statistics, classical positions are not able to fully represent their way of playing, and a new set of 5 roles emerges as a meaningful classification of players' characteristics.

keywords: Sport Analytics, Self-Organizing Maps, Fuzzy Clustering, Basketball.

1 Introduction

Advanced analytics in modern sports can help the coaching staff and management to face different problems. They can be especially effective to address the assessment of the

*Corresponding author: tullio.facchinetti@unipv.it.

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

performance of teams and single players. Problems can be related to any aspect of the game and coaching activity, from planning optimal training sessions in different parts of the season, to deciding the best strategy to adopt against the next opponent team, or to improve some specific aspects of the team game.

One major everyday problem to face is the selection of the best set of players to line up on the field for achieving the best result at any time during the game. This selection is a complex and delicate task for a coach, since the decision can be influenced by many factors.

Most team sports are characterized by positions to which the players in a team are classified, to assign them specific responsibilities during a game. In many cases, positions have been defined with the appearance of the related sport, and many sports were born a long time ago. There are several examples of positions that were defined at the origin of a given sport, and are nowadays no longer useful to describe the way players play the game. In some cases, it happens that players are changing the way of interpreting the meaning of their position. Although in most cases it is difficult to find a radical change of all the positions in a specific game, there are many examples of great players in the history who really changed the way of interpreting a position, or even created a new one thanks to their way of playing the game.

The case of the *mobile quarterback* is discussed in Monson (2012). This dissertation describes how some elite athletes, with excellent technical passing skills and game vision, have redefined the characteristics of the quarterback in the US National Football League, allowing many more options in the game w.r.t. the traditional gameplay. Corsello (2016) discusses how *Manuel Neuer*, one of the best soccer goalkeepers nowadays, is going to potentially redefine the goalkeeper position. His peculiar, innovative (and risky) ability to play far from the goal, with excellent technical skills, extends the area when the goalkeeper becomes effective. Mahmood (2015) explains how *the Makelele position* is deeply affecting the English football (soccer) team organization. Makelele, a midfielder featuring athleticism, spirit of sacrifice and defensive attitude, was mainly devoted to interfere with the opponents attack. In this way, Makelele revealed to be fundamental in important teams with very high offensive attitude, giving balance to the gameplay and performing as a defensive coordinator from the midfield. Finally, Flannery (2016) discusses how *Draymond Green* is redefining the concept of NBA basketball superstar. Being an hybrid frontcourt player[1], he is outstanding in every position, making it difficult to label him with a traditional category.

The aforementioned works consider outstanding players, who are changing how to play a specific position in their sport. However, none of them address a complete revolution or redefinition of all the positions in a sport, which is the ultimate goal of the present work. Moreover, they tackle the problem from a non-scientific perspective, but leverage the opinions and the analysis of domain experts.

The focus of this paper is on basketball. In general, the 5 players on the basketball court are selected according to their positions, so that the coach selects one player from

[1]A frontcourt in basketball is the player who usually plays the power forward or center position; often is one of the tallest players in the team.

Table 1: The traditional positions of basketball.

Position	Characteristics
Point Guard or Playmaker	He is intended to be the "brain" of the team, the player who "creates the game"; the Point Guard is often the shortest player in the team.
Shooting Guard	Usually the main duty of a Shooting Guard is to score points far from the basket; there are exceptions represented by defensive minded players, whose role in the team is to stop the best offensive player of the opposite team.
Small Forward	Small Forwards are usually very athletic players who can score from different areas of the field; they also help in defending and rebounding.
Power Forward	Power Forwards are expected to play closer to the basket, mainly helping the team by scoring points and taking rebounds.
Center	The Center is supposed to be the biggest player in the team, the one who grabs a lot of rebounds, protects the rim on defense, blocking shots and taking advantage of his size on offense.

each one of the 5 available positions. Since Mr. James Naismith invented basketball in 1891, the same 5 positions have been used to identify the role of players composing a team. Table 1 reports the list of the 5 available positions and their traditionally associated characteristics. The reported positions are not only labels assigned to players to help selecting the 5 ones that will start the game or will be on the court at the same time, but they have always been associated to specific characteristics of the players associated to the position.

An interesting problem that raised during the last few years regarding basketball analysis is the adaptation of player's positions to the modern basketball. In fact, the basketball changed quite a lot since its appearance, and it is nowadays rather different from the game invented by Mr. Naismith for his students. One of the most disrupting change in the game was the introduction of the 3 point line in the early 80's, which deeply changed the way of playing the game all over the world (Sanders (1981); Butts (1986); Lynch (1987)). The availability of a new kind of shot, rewarded with one extra point, obviously increased the interest at shooting from a higher distance. The new rule raised the importance of long shooting skills, giving more scoring responsibilities to players who were playing far from the basket, and changing the way of playing in such positions.

The evolution of the game led to the consequent evolution of players, who became more athletic, more skilled, and developed different ways to play the game. Due to such evolution, it may be difficult in some cases to assign a player to one of the 5 positions that have been traditionally used in the history of basketball. It is common to find players that are halfway between two classic positions, or players that are labeled with the same position but actually play the game in very different ways.

A seminal work on basketball player's position characterization was presented in Alagappan (2012). The focus is on 2 main problems of the traditional basketball positions: *oversimplification* and *incorrect classification*. Oversimplification is due to the fact that the traditional 5 positions are failing at expressing the diversity of the game, since nowadays there are tens of different ways of playing basketball world-wide. On the other hand, the incorrect classification of modern players into a given position is due to subjective factors. The author addresses the problem of making this classification more objective. The research applies the Topological Data Analysis (TDA, see Carlsson (2009)) to NBA players statistics to show differences between players statistical profiles and to redefine basketball positions according to the results of the analysis. The goal is to show the differences between data samples through shape analysis. The TDA is applied to players statistics to find differences between players and the way they play the game. This approach allows to obtain a *similarity network* where data points represent some elements of the considered dataset, while edges connecting nodes capture the similarity between nodes.

The method was applied to NBA players of the 2010-11 NBA regular season. Each node in the network represents one or more NBA players. The players are mapped on the network according to the similarity of their statistics. The parameters used in the analysis were points, rebounds, assists, steals, turnovers, fouls and blocks. The values were normalized according to the minutes played per game (MPG).

The result was a set of 13 clusters, corresponding to 13 new positions unveiled in the analysis. The TDA allows to compare players that are clustered in a certain position and to analyze which ones are closer to the prototype of a player playing in a certain position. For instance, Jason Terry and Tony Parker result to be the closest player to the perfect Offensive Ball-Handler[2], while Marcus Camby and Tyson Chandler result very good examples of Paint Protectors[3].

Beside showing the new positions derived from the map and how these positions identify the way players play the game, the author mapped the spatial formation of players from specific teams of the 2010-2011 NBA season, assigning the players of these teams to the new positions. The proposed new positions were used to analyze the composition of different teams. This allowed to put into relationship the composition of a team with winning and losing records, and to investigate the existence of common patterns of players positions that influence the performance of a team. In other words, the identification of a combination of players that is common to all the teams with good performance and winning rate would allow to recreate a team that is more likely to be successful.

[2] Jason Terry was a Shooting Guard for the Dallas Mavericks, while Tony Parker was the starting Point Guard of the San Antonio Spurs.

[3] Marcus Camby and Tyson Chandler are both centers. Camby was playing for the Portland Trail-Blazers, while Chandler was playing for the Dallas Mavericks.

Alagappan remarks that the possibility to truly describe the features of a player might help general managers to organize more balanced teams and to reliably determine whether a player fits in the team, just looking at his *true position.* A suitable interpretation of the positions can also help the coach when he builds a lineup or evaluates offensive and defensive strategies. Positions that better describe the way players interpret the game allow the coach to create strategies that can better combine the characteristics of players, leveraging their skills at the maximum potential.

Alagappan (2012) suggests that a meaningful change in the way of interpreting basketball is required to avoid remaining stuck to obsolete conventions. Moreover, it shows that the evolution of the basketball as a game makes the traditional 5 positions no longer suitable for today's needs of coaches, general managers and enthusiasts.

The present paper takes as a main source of inspiration the ideas in Alagappan (2012) and investigates NBA data using machine learning and data mining procedures. As a first result, we confirm the inadequateness of classical positions to the nowadays way of playing and, ultimately, the idea of *incorrect classification* pursued by Alagappan. As far as it is concerned with the second Alagappan's idea, *oversimplification*, we follow a different approach: instead of defining a high number of clusters, we limit ourselves to design few new roles, but we use a fuzzy clustering procedure that, as a matter of fact, describes each player with a unique combination of his memberships to the rough positions. The goal is to define a novel way of interpreting modern basketball, to help professionals, experts and fans to better understand the game and to enable deeper analyses of results.

New positions are proposed on the basis of an analytic and objective method, considering the statistical profile of players. The analysis starts from the statistical profile of a player during a season. Players with similar statistical profiles will fit in the same newly defined position, while players with different statistical profiles will belong to different positions.

To define the new positions, we propose an approach based on the joint use of fuzzy clustering and Self-Organizing Maps (SOM) (Kohonen, 1982), a class of unsupervised neural networks that applies a competitive learning rule (as opposed to error-correction learning, such as for example backpropagation) based on a neighborhood function aimed at preserving the topological properties of the input space. The SOM is fed with selected statistical features of a set of NBA players. Afterwards, a fuzzy clusterization of players is obtained by means of a polynomial fuzzifier function applied to the output layer of the SOM. A basketball technical analysis of the cluster profiles, together with the consideration of the membership coefficients of the players to the different clusters, will finally allow to classify all the players into a number of groups, corresponding to the newly defined positions. The paper is organized according the structure Introduction (Sections 1 and 2), Methods (Section 3), Results (Section 4), and Discussion (Section 5).

Table 2: Atypical players statistics

PLAYER	TRB	BLK	AST	STL	TOV	PF	PPG
LeBron James	7.4	0.6	6.8	1.4	3.3	1.9	25.3
Kawhi Leonard	6.8	1.0	2.6	1.8	1.5	1.8	21.2
Russell Westbrook	7.8	0.3	10.4	2.0	4.3	2.5	23.5
Draymond Green	9.5	1.4	7.4	1.5	3.2	3.0	14.0

2 Classic positions analysis

This section analyzes the traditional basketball positions to show how they are not able to describe modern players game anymore.

The training set used in this analysis was composed of 5 elements, one for each position: Point Guard, Shooting Guard, Small Forward, Power Forward and Center. Each element of the training set is obtained as the season average statistics of the players of the same position. We decided to use the same 7 statistics used by Alagappan (2012), *i.e.* Average Total Rebounds per Game (TRB), Average Blocks per Game (BLK), Average Assists per Game (AST), Average Steals per Game (STL), Average Turnovers per Game (TOV), Average Personal Fouls per Game (PF), Average Points Per Game (PPG). Since these are just a subset of the several available statistics describing the way of playing, the whole analysis has been repeated using a larger subset of 11 statistics (where TRB has been decomposed into Offensive and Defensive Rebounds, and three statistics have been added, concerning the Average Shooting Percentages for 2-point, 3-point shots and free throws). The results do not appreciably differ from those obtained with Alagappan's 7 statistics (Bianchi, 2016), that will be presented hereafter. After the training process, the output layer of the SOM was divided into 5 clusters, each one corresponding to one position. The color assigned to a neuron is selected according to its similarity with the elements of the training set. Since each element of the training set is a reference to a particular position, the color of the neuron will be one of the 5 colors assigned to the 5 positions.

The center of each cluster obtained on the output layer is the position where the elements of the training set were mapped. Therefore, the forward mapping process feeding the SOM with the statistical profiles of some particular players allows to obtain the place in the map corresponding to the 5 classic positions.

Firstly, some of the most atypical players in the league were mapped on the SOM. The position of these players have been largely debated during the last seasons. In fact, although they are labeled with a certain classic position, their way of playing that position is different from all the other players in the league. The selected players are LeBron James, Kawhi Leonard, Draymond Green and Russell Westbrook. Their statistics are reported in Table 2.

LeBron James is considered to be a small forward. Actually, he is taller and bigger than most of the other small forwards in the league, he can score as a shooting guard,

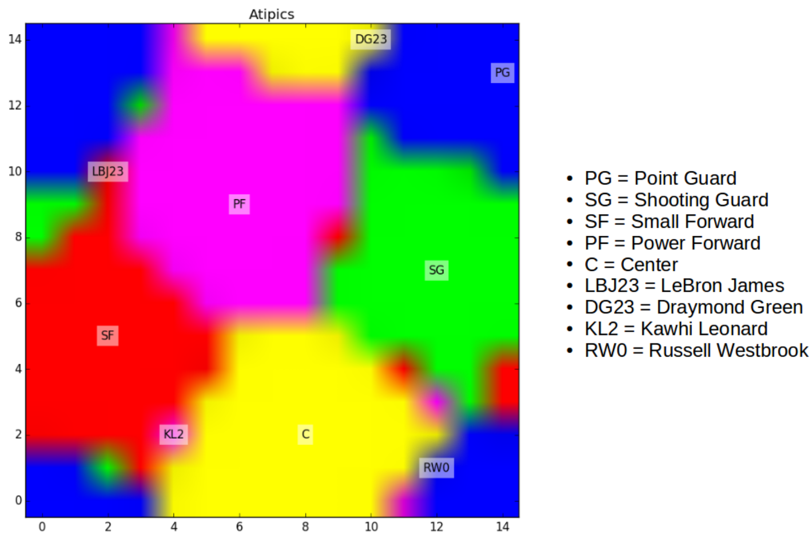

Figure 1: Forward mapping of some atypical players on the 5 positions trained SOM

pass the ball as a point guard, and grab rebounds like a power forward. He played all of these positions during his career. Kawhi Leonard is considered a small forward too, but his rebounding, scoring and defensive skills, make him a unique player. Draymond Green is probably the most atypical player, since he can do almost everything on a basketball court and he was defined as the epitome of the position-less basketball player. Finally, Russell Westbrook is labelled as a point guard. In terms of statistics, he is one of the best athlete who ever played the position in basketball history. He made 18 triple doubles[4] during the 2015-2016 regular season, tying the previous record made by Magic Johnson in the 1981-1982 regular season. A point guard is expected to be a good scorer and assist-man, but not necessarily a great rebounder or shot blocker. Westbrook has these skills altogether.

From Figure 1, which shows the result of the forward mapping of these atypical players on the output layer of the SOM, it is immediately clear how classic positions are not suitable to describe these players. LeBron James (LBJ23) is mapped in the middle among small forward players, power forwards, point guards and shooting guards. He is half-way among all the positions previously cited to describe his way of playing. Kawhi Leonard (KL2) is mapped between the small forwards and centers; he is considered a small forward, but his rebounding and defensive skills place him close to centers. Russell Westbrook (RW0) and Draymond Green (DG23) are in the peripheral area of the point guard position, but they are on the edge which is close to the center position. For the way classic basketball positions are intended, point guards and centers should be two opposite types of players, with completely different responsibilities on the floor, and so

[4] A triple double happens when a player records a value that is greater or equal to 10 in three different statistics during one game. It means for example that in one game a player has to score at least 10 points, grab at least 10 rebounds and give at least 10 assists to make a triple double.

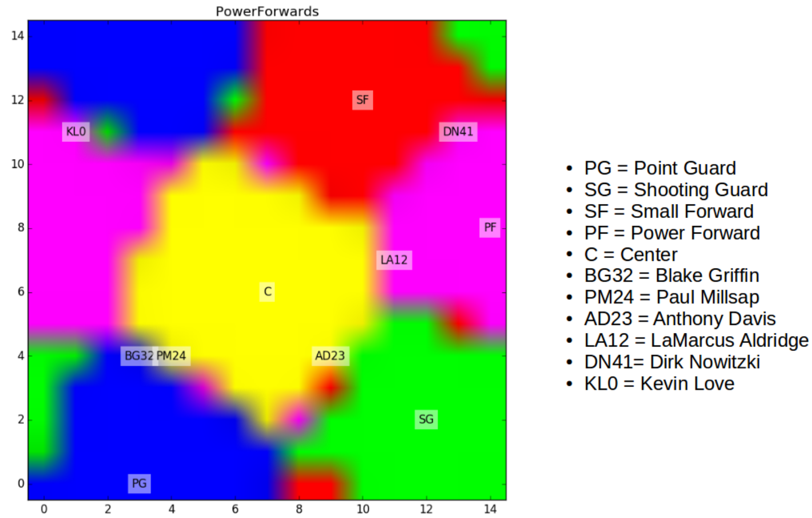

Figure 2: Forward mapping of some power forwards on the 5 positions trained SOM

Table 3: Power Forwards statistics

PLAYER	TRB	BLK	AST	STL	TOV	PF	PPG
Kevin Love	9.9	0.5	2.4	0.8	1.8	2.1	16.0
Paul Millsap	9.0	1.7	3.3	1.8	2.4	2.9	17.1
Blake Griffin	8.4	0.5	4.9	0.8	2.4	2.7	21.4
Anthony Davis	10.3	2.0	1.9	1.3	2.0	2.4	24.3
LaMarcus Aldridge	8.5	1.1	1.5	0.5	1.3	2.0	18.0
Dirk Nowitzki	6.5	0.7	1.8	0.7	1.1	2.1	18.3

completely different ways of playing the game, as explained in Section 1. The fact that these players are mapped half-way between point-guards and centers shows how the classic positions are not able to suitably describe their way of playing.

A second experiment was made on the SOM clustered considering the 5 classic positions by selecting some players labeled to be playing the same positions, and forward-mapping them on the output layer. The goal is to show how the way of playing the same position can be rather different from player to player. The first test was made by mapping some of the most famous power forwards in the league: Kevin Love, Paul Millsap, Blake Griffin, Anthony Davis, LaMarcus Aldridge and Dirk Nowitzki, whose statistics are reported in Table 3. Figure 2 shows their positions on the output layer on the SOM. These players are not mapped in the center of the purple area related to power forwards, but are spread across different positions in the map, even if most of the them are on the edge of the power forward area.

The most interesting observation is regarding their relative distance, since it is the

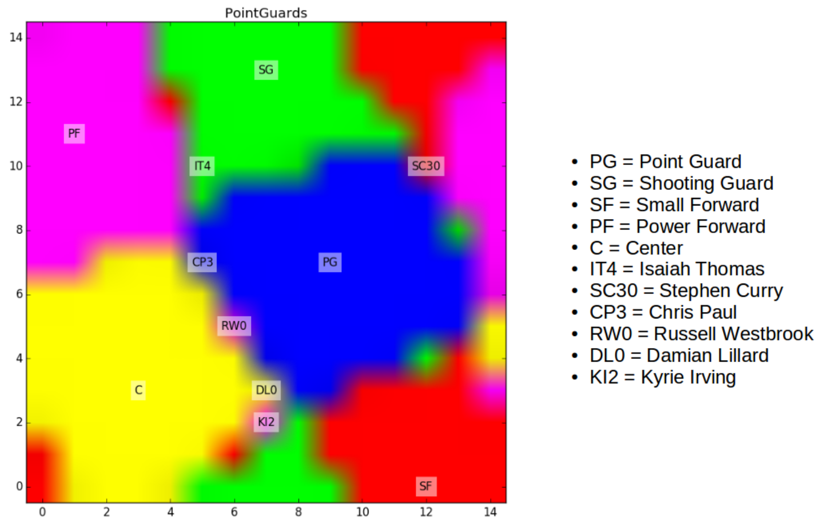

Figure 3: Forward mapping of some point guards on the 5 positions trained SOM

parameter that actually describes the difference between their way of playing. Paul Millsap (PM24) and Blake Griffin (BG32) are mapped in very close positions, since their TRB, TOV and PF statistics have similar values. Moreover, they are the two players in the set with the highest values in AST. Anthony Davis (AD23), with his high values of TRB, BLK and PPG, is mapped in a position that is half-way between the center and the shooting guard positions, two roles that should have almost nothing in common, but Davis can be considered as a player who grabs rebounds and blocks shots like a center and scores like a shooting guard, making the power forward position unsuitable to fit his game. LaMarcus Aldridge (LA12), Dirk Nowitzki (DN41) and Kevin Love (KL0) are the 3 players who are mapped close to the power forward cluster, but they are very distant from each other. Love is a good scorer, but his game is characterized by an high number of TRB and a good value of AST. The good rebounding skills make him a good power forward, but his passing skills place him in a position of the map which is close to the point guard area, like in the case of Millsap and Griffin. Aldridge and Nowitzki instead have statistics that are characterized by good scoring and rebounding numbers, but no other excellent statistics. They are both placed on the edge of the power forward position, but Aldridge, who is a better rebounder, is close to the center region, since he has higher TRB and BLK statistics, while Nowitzki, who is a poorer rebounder and shot blocker, is placed close to the small forward area.

The same type of experiment was carried out on the point guard position, mapping some of the most famous point guards in the league on the output layer of the SOM.

The players selected in this test are Kyrie Irving, Isaiah Thomas, Damian Lillard, Stephen Curry, Chris Paul and Russell Westbrook, whose statistics are reported in Table 4. Figure 3 shows that they are less spread w.r.t. power forwards, since they are closer to the point guard blue region of the map. Even if they are close to the point guard region, their relative distances show that there are some differences between their

Table 4: Point Guards statistics

PLAYER	TRB	BLK	AST	STL	TOV	PF	PPG
Kyrie Irving	3.0	0.3	4.7	1.1	2.3	2.0	19.6
Isaiah Thomas	3.0	0.1	6.2	1.1	2.7	2.0	22.2
Damian Lillard	4.0	0.4	6.8	0.9	3.2	2.2	25.1
Stephen Curry	5.4	0.2	6.7	2.1	3.3	2.0	30.1
Chris Paul	4.2	0.2	10.0	2.1	2.6	2.5	19.5
Russell Westbrook	7.8	0.3	10.4	2.0	4.3	2.5	23.5

way of playing the position. Russell Westbrook (RW0) and Chris Paul (CP3), for example, are placed in close positions, between the point guard area, the power forward area and the center area. This is due to the fact that they are the two players with the highest values of AST and they have very similar statistics in STL and PF. They are both good scorers, but Westbrook is better in TRB, a feature that maps him closer to the center yellow region, while still being classified as power forward. Kyrie Irving (KI2), Damian Lillard (DL0) and Isaiah Thomas (IT4) have similar statistical profiles, having low values of TRB, and STL and AST w.r.t. the other players in the set, in particular Westbrook and Paul, while having average values of PPG. This differences bring them in the farthest positions w.r.t. the point guard area, since Thomas and Irving are closer to the shooting guard green area. Lillard is placed half-way between the blue point guard, the green shooting guard and the yellow center areas. The main reason for being closer to the center area is that he is the point guard with the higher value in BLK in the set, he has also good rebounding ability compared to the other players in the set, even if not excellent, and an high number of AST. This good performance in several statistics makes it harder to map him into a well defined cluster.

Finally, Stephen Curry (SC30) is mapped on the opposite side of the point guard cluster w.r.t. other players, closer to the small forward and the shooting guard clusters. He has high values in all the categories, being the best, or close to the best player in the set according to each category. He is the best scorer, with 30.1 PPG, a value that is much higher than the others[5]; moreover, he has very good numbers in AST, TRB and STL with respect to the other players. With high numbers in almost all the categories, Curry is an outstanding player, and his position is half-way between the blue point guard, the red small forward, the green shooting guard and the purple power forward areas, making him difficult to be assigned to a precise position too.

The last two cases show that there are different ways of playing the same position. The statistics can highlight the difference between players, who are thus required to be classified into different categories. The results show that the classic basketball positions are not able to fit the way many modern players play the game anymore, and in many cases it is not possible to completely describe a player with just one position. The results

[5]In the 2015-2016 regular season Stephen Curry has been the player with the highest PPG in the whole NBA.

of these experiments represent the motivation of our goal to provide new positions and new criteria to classify basketball players.

3 Methods

In this section the data used for the analysis (Subsection 3.1) and the procedure followed in order to define the new positions (Subsection 3.2) are described. The technical discussion of results will be presented in Section 5. With regard to the first point, it is worth to outline that the training set is crucial when using a neural network, and the SOM is not an exception. Indeed, the training set used to train the network affects the clusters that will appear in the output layer after the training. These clusters will be used later to assign some elements of the input space to the corresponding categories. For this reason, during the experiments, different training sets have been used to explore different results of the SOM clustering. This section describes the training set that allowed to obtain the most interesting results. Details about the other analyses can be found in Bianchi (2016).

3.1 Key Players Training Set

The main source of data used for the experimental validation of this work was the NBA database of statistics. In particular, the experiments discussed in this paper were based on data referring to the 2015-2016 NBA regular season. The season includes 82 games played by the 30 teams in the league, involving 476 players. The players are described according to the same 7 statistics used for the analysis presented in Section 2: TRB, BLK, AST, STL, TOV, PF, PPG.

One possible option about the training set is to select the 5 starting players of each team. However, many teams in the NBA do not build their success only on the players belonging to the starting five, but also on the fundamental contribution from players who start a game from the bench. For this reason, a more adequate training set can be composed by the 8 players with the highest number of average minutes played per game, for each of the 30 teams (Table 5). Although the five starting players are usually the ones that average the highest number of minutes played, in some cases the players coming from the bench may play more minutes than some starters, thus providing a very important help to the team. Therefore, this training set, hereafter denoted as *Key Players Training Set*, better reflects the true core of each team, made by the five starters and some key players coming from the bench.

3.2 Classification procedure

The procedure was carried out following three main steps:

1. definition of a 7-dimensional SOM trained with the Key Players Training Set;

Table 5: Key Players Training Set (in parenthesis: average minutes played per game)

TEAM	PLAYERS								AVG_MP
ATL	P. Millsap (32.7)	A. Horford (32.1)	K. Korver (30.0)	J. Teague (28.5)	K. Bazemore (27.8)	T. Sefolosha (23.4)	D. Schroder (20.3)	T. Hardaway (16.9)	26.46
BOS	A. Bradley (33.4)	I. Thomas (32.2)	J. Crowder (31.6)	E. Turner (28.0)	M. Smart (27.3)	J. Sullinger (23.6)	A. Johnson (22.8)	K. Olynyk (20.2)	27.39
BRK	B. Lopez (33.7)	J. Johnson (33.4)	T. Young (33.0)	J. Jack (32.1)	B. Bogdanovic (26.8)	S. Larkin (22.4)	D. Sloan (21.6)	W. Ellington (21.3)	28.04
CHI	J. Butler (36.9)	P. Gasol (31.8)	D. Rose (31.8)	T. Gibson (26.5)	N. Mirotic (24.9)	D. McDermott (23.0)	M. Dunleavy (22.7)	J. Noah (21.9)	27.44
CHO	K. Walker (35.6)	N. Batum (35.0)	C. Lee (29.5)	M. Kidd-Gilchrist (29.3)	M. Williams (28.9)	J. Lin (26.3)	C. Zeller (24.3)	A. Jefferson (23.3)	29.03
CLE	L. James (35.6)	K. Irving (31.5)	K. Love (31.5)	J. Smith (30.7)	T. Thompson (27.7)	M. Dellavedova (24.6)	I. Shumpert (24.4)	M. Williams (18.2)	28.02
DAL	W. Matthews (33.9)	D. Williams (32.4)	D. Nowitzki (31.5)	C. Parsons (29.5)	R. Felton (27.4)	Z. Pachulia (26.4)	J. Barea (22.5)	D. Harris (20.0)	27.95
DEN	D. Gallinari (34.7)	G. Harris (32.1)	E. Mudiay (30.4)	W. Barton (28.7)	J. Nelson (26.6)	K. Faried (25.3)	N. Jokic (21.7)	D. Arthur (21.7)	27.65
DET	K. Caldwell-Pope (36.7)	M. Morris (35.7)	T. Harris (33.1)	A. Drummond (32.9)	R. Jackson (30.7)	E. Ilyasova (25.4)	S. Johnson (23.1)	A. Tolliver (18.6)	29.52
GSW	D. Green (34.7)	S. Curry (34.2)	K. Thompson (33.3)	H. Barnes (30.9)	A. Iguodala (26.6)	A. Bogut (20.7)	S. Livingston (19.5)	F. Ezeli (16.7)	27.07
HOU	J. Harden (38.1)	T. Ariza (35.3)	D. Howard (32.1)	P. Beverley (28.7)	T. Jones (20.9)	C. Brewer (20.4)	C. Capela (19.1)	M. Beasley (18.2)	26.60
IND	P. George (34.8)	G. Hill (34.1)	M. Ellis (33.8)	I. Mahinmi (25.6)	C. Miles (22.9)	M. Turner (22.8)	R. Stuckey (22.0)	T. Lawson (21.4)	27.18
LAC	D. Jordan (33.7)	B. Griffin (33.4)	C. Paul (32.7)	J. Green (28.2)	J. Redick (28.0)	J. Crawford (26.9)	A. Rivers (21.9)	W. Johnson (20.8)	28.20
LAL	J. Clarkson (32.3)	L. Williams (28.5)	K. Bryant (28.2)	D. Russell (28.2)	J. Randle (28.2)	R. Hibbert (23.2)	A. Brown (20.7)	B. Bass (20.3)	26.20
MEM	M. Gasol (34.4)	M. Conley (31.4)	Z. Randolph (29.6)	M. Barnes (28.8)	T. Allen (25.3)	J. Farmar (24.3)	B. Weber (24.1)	M. Chalmers (22.5)	27.55
MIA	C. Bosh (33.5)	G. Dragic (32.8)	L. Deng (32.4)	D. Wade (30.5)	H. Whiteside (29.1)	J. Winslow (28.6)	T. Johnson (24.0)	G. Green (22.6)	29.19
MIL	K. Middleton (36.1)	G. Antetokounmpo (35.3)	J. Parker (31.7)	M. Carter-Williams (30.5)	G. Monroe (29.3)	J. Bayless (28.9)	O. Mayo (26.6)	G. Vasquez (20.0)	29.80
MIN	A. Wiggins (35.1)	K. Towns (32.0)	R. Rubio (30.6)	Z. LaVine (28.0)	G. Dieng (27.1)	S. Muhammad (20.5)	T. Prince (19.0)	N. Bjelica (17.9)	26.27
NOP	A. Davis (35.5)	E. Gordon (32.9)	T. Evans (30.6)	R. Anderson (30.4)	J. Holiday (28.2)	J. Hamilton (27.6)	N. Cole (26.6)	D. Cunningham (24.6)	29.55
NYK	C. Anthony (35.1)	A. Afflalo (33.4)	K. Porzingis (28.4)	J. Calderon (28.1)	R. Lopez (27.1)	L. Galloway (24.8)	L. Thomas (22.3)	D. Williams (17.9)	27.14
OKC	K. Durant (35.8)	R. Westbrook (34.4)	S. Ibaka (32.1)	D. Waiters (27.6)	S. Adams (25.2)	A. Roberson (22.2)	E. Kanter (21.0)	R. Foye (20.3)	27.32
ORL	V. Oladipo (33.0)	E. Fournier (32.5)	N. Vucevic (31.3)	E. Payton (29.4)	A. Gordon (23.9)	C. Watson (19.9)	B. Jennings (18.1)	M. Hezonja (17.9)	25.75
PHI	J. Okafor (30.0)	N. Noel (29.3)	I. Smith (29.1)	R. Covington (28.4)	H. Thompson (28.0)	J. Grant (26.8)	I. Canaan (25.5)	N. Stauskas (24.8)	27.74
PHO	B. Knight (36.0)	E. Bledsoe (34.2)	P. Tucker (31.0)	D. Booker (27.7)	T. Chandler (24.5)	A. Len (23.3)	T. Warren (22.8)	M. Teletovic (21.3)	27.60
POR	D. Lillard (35.7)	C. McCollum (34.8)	A. Aminu (28.5)	A. Crabbe (26.0)	M. Plumlee (25.4)	M. Leonard (21.9)	E. Davis (20.8)	G. Henderson (19.9)	26.63
SAC	R. Rondo (35.2)	D. Cousins (34.6)	R. Gay (34.0)	D. Collison (30.0)	O. Casspi (27.2)	M. Belinelli (24.6)	D. Dukan (24.0)	W. Cauley-Stein (21.4)	28.88
SAS	K. Leonard (33.1)	L. Aldridge (30.6)	T. Parker (27.5)	D. Green (26.1)	T. Duncan (25.2)	P. Mills (20.5)	K. Martin (19.9)	M. Ginobili (19.6)	25.31
TOR	K. Lowry (37.0)	D. DeRozan (35.9)	D. Carroll (30.2)	J. Valanciunas (26.0)	C. Joseph (25.6)	P. Patterson (25.6)	T. Ross (23.9)	B. Biyombo (22.0)	28.28
UTA	G. Hayward (36.2)	R. Hood (32.2)	D. Favors (32.0)	R. Gobert (31.7)	A. Burks (25.7)	T. Burke (21.3)	T. Booker (20.7)	S. Mack (20.3)	27.51
WAS	J. Wall (36.2)	B. Beal (31.1)	O. Porter (30.3)	M. Gortat (30.1)	J. Dudley (25.9)	M. Morris (25.5)	G. Temple (24.4)	R. Sessions (20.3)	27.98

2. clusterization of the SOM output layer into a proper number of groups by means of a fuzzy clustering algorithm, to take into account the possible presence of players whose statistics' values suggest a hybrid position;

3. analysis of the clusters' profiles to detect heterogeneities that allow to define more refined positions within the clusters, resorting to basketball technical considerations.

With regard to step 1, it is worth to remember that no particular clustering or shape of the output layer was expected or forced during the experiment. In fact, the aim of the SOM is to automatically model the shape of the output layer, on the basis of the data characterizing the input space, depending on the training set used to feed the network during the learning process. The number of neuron of the output layer is set to 900, so a 900×7 data matrix was passed to the second step for fuzzy clustering.

In step 2 the fuzzy k-means clustering algorithm with polynomial fuzzifier function (Winkler et al., 2011) was applied. The polynomial fuzzifier allows to obtain areas of crisp membership degrees around the cluster centroids, while membership coefficients are computed outside of these areas. Therefore, membership coefficients are equal to 1 for objects very close to the cluster centroids (thus definitely assigned to clusters), while objects with hybrid positioning are clearly detected. To decide the number of clusters, 4 indexes were computed: the Partition Coefficient (Bezdek, 1973), the Partition Entropy (Bezdek, 2013), the Silhouette (Kaufman and Rousseeuw, 2009), and the Fuzzy Silhouette (Campello and Hruschka, 2006). They all suggested $k = 3$ clusters as the optimal choice (Figure 4). This clusterization recognizes the presence of 3 main agglomerations of the output layer neurons in the 7-dimensional space. However, wide areas of fuzziness

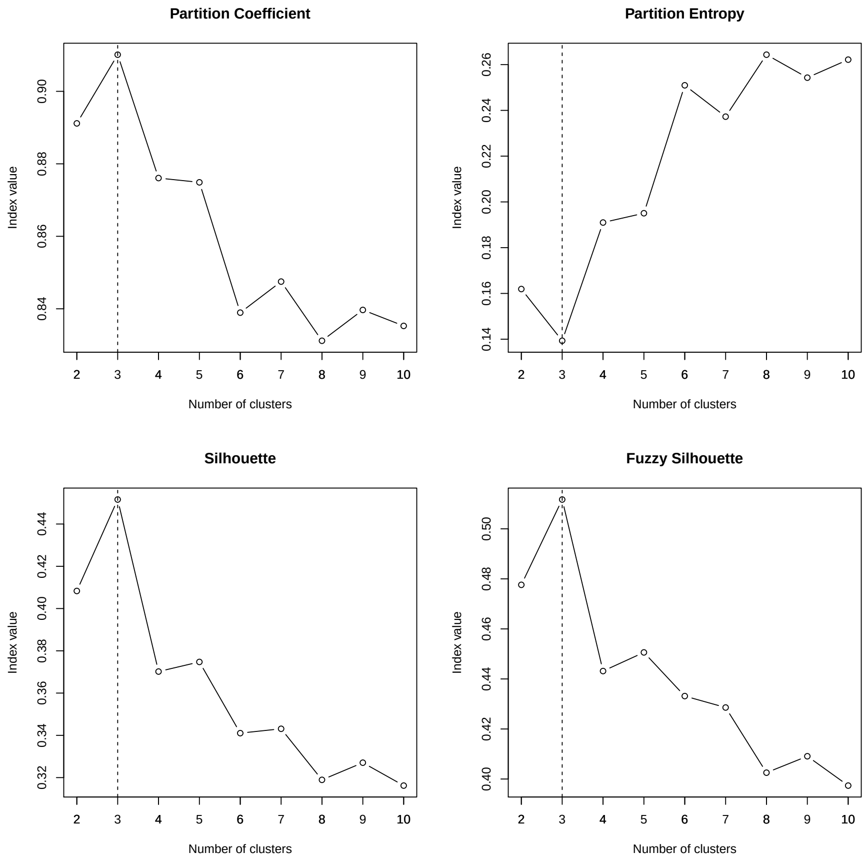

Figure 4: Indexes for the choice of the optimal number of clusters.

are also defined. The graph in Figure 5 reports the boxplots of the membership coefficients, ranging from about 0.33 to about 0.92, with no appreciable differences among clusters. Moreover, we have to point out that, from a basketball technical point of view, the definition of only three positions would not be an admissible solution. These remarks highlight the need for a finer analysis within the clusters, aimed at examining the particular profiles of players with hybrid position and, possibly, the definition of subgroups with similar features (step 3).

In step 3, to visually inspect the positioning of the players in the output layer of the SOM, a dimensionality reduction was performed by means of MultiDimensional Scaling (MDS). The resulting 2-dimensional representation is displayed in Figure 6, where the clusters obtained in step 2 are delimited by blue lines. Combining both statistical and basketball considerations, a finer analysis of the players included in each cluster, together with their membership coefficients, led to recognizing some subgroups within the main clusters, delimited by red lines. The features of these subgroups, the reasons for their definition, and the evaluation of some example cases are discussed in Section 5.

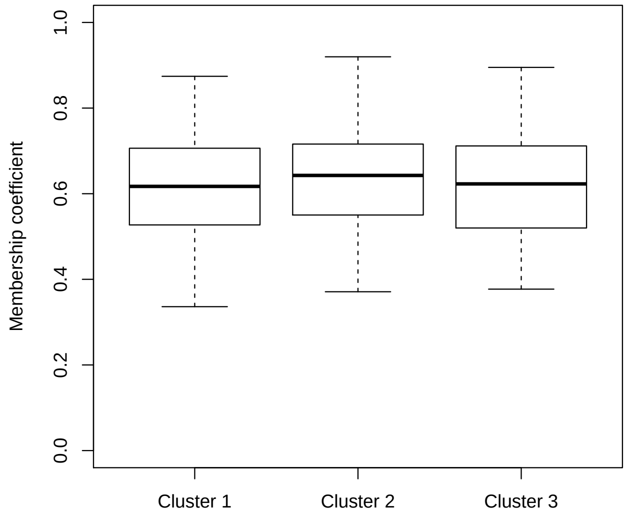

Figure 5: Boxplots of the membership coefficients for the three clusters.

4 Results

As mentioned in Section 3, the new positions have been obtained by combining statistical and basketball considerations. In fact, after three main clusters of players were defined by means of the fuzzy clustering algorithm, we resorted to basketball technical considerations to recognize finer subgroups within the clusters.

As a result, the analysis leads to the identification of the new positions as follows:

1. **All-Around All Star** (AAS - Cluster 1). A player belonging to this category is outstanding in many aspects of the game, with statistics that are significantly higher than the league averages in most of the statistical parameters. AAS players can make the difference for their teams in many different ways. They are usually elite scorers, but combine the scoring ability with great passing skills; alternatively, they may have big rebounding numbers or great defensive skills; in same cases all these characteristics are present at the same time.

2. **Scoring Backcourt** (SB - Cluster 1). SB players are characterized by remarkable offensive skills. These players have high values of PPG statistics, but usually they also have good passing skills, which means high values in AST. They are usually below the average in statistical figures such as TRB and BLK, since they perform better when they play away from the basket.

3. **Scoring Rebounder** (SR - Cluster 2). A SR is a player with high average statistics mainly in PPG and TRB. These players are skilled scorers, both in low-post

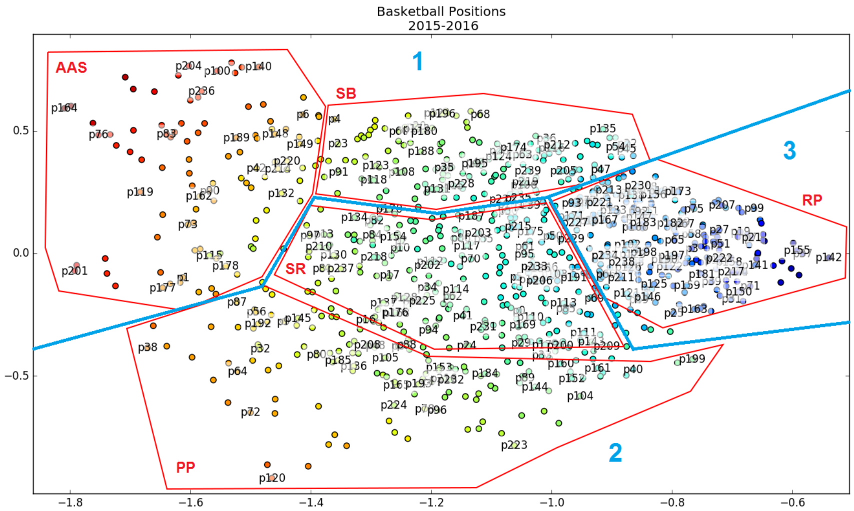

Figure 6: Two-dimensional representation via MDS of the SOM output layer, with fuzzy clusterization and subgroups.

game[6] and facing the basket, with special rebounding abilities. Their size is bigger w.r.t. SB and are used to play closer to the basket. Thanks to these peculiar technical and athletic skills, they take advantage of the position on the field and their size to score from a very short distance.

4. **Paint Protector** (PP - Cluster 2). A PP player is usually not a great scorer, with PPG values that are below the averages of the league. They are very good at rebounding and blocking shots, thanks to their size and excellent defensive skills As a result, their values related to TRB, BLK and STL are higher than the other statistical figures. They usually perform a more athletic game and play closer to the basket on both ends of the floor.

5. **Role Player** (RP - Cluster 3). The RP is usually very good but not excellent in only one statistical category[7]; in some cases, he has numbers that are slightly below the average of the league in all statistical figures. This group includes players who are considered as specialists of particular aspects of the game. They can be outside shooters, whose main responsibility is to finalize the actions of other players, or defensive specialists, whose task is to guard the best player of the opponent team. These roles can be crucial in a team, but their importance is usually not reflected in their statistics, which are usually lower than other players. This is probably the reason why those players are clustered together on the output layer of the SOM.

[6] A low-post scorer is a player who is able to score points playing closer to the basket, i.e., the so-called low-post position. Playing in the low-post means being able to receive the ball close to the basket, giving the back to it.

[7] Most of the times when a RP is very good but not excellent in just one statistical category, this category is PPG.

The result of the clustering is shown in Figure 6. The result shows 5 red highlighted areas in the figure, which identify the mapping of players into the 5 new positions.

The 5 areas and positions have been derived from the 3 clusters that represent the outcome of the fuzzy analysis. Clusters 1 and 2 have been further refined in 2 sub-clusters each, corresponding to two additional positions, on the base of the following considerations.

Cluster 1 contains a set of players who have a very good scoring and passing game; this fact brings them above the average in PPG and AST statistical categories. Players who are assigned to All-Around All Star position (AAS) have slightly higher statistics in PPG and AST w.r.t. Scoring Backcourt players. However, what makes the difference between those two categories are the numbers in all the other statistics: AAS players have better figures in TRB, BLK and STL w.r.t. SB colleagues. As a result, AAS players have stronger ability to affect the overall game, since they contribute to every aspect of the game itself, helping their team in many different ways.

Cluster 2 instead groups players who have in common the ability to impact the game close to the basket, even if in different ways. Both Scoring Rebounder and Paint Protector players have TRB and BLK statistics that are above the average of the rest of the league. Usually PP players are excellent in those two statistical categories, but SR players are much better in PPG and AST, even if not as good as players included in Cluster 1. This distinction is very important in practice, and it is the main reason why Cluster 2 has been divided to reflect this situation. PP players do usually affect the game in specific ways. They take advantage of their physical game, helping the team rebounding and blocking shots, but usually they have poor offensive responsibilities during the game. On the other hand, SR players are more versatile and, apart from this defensive effort, they usually bring more offensive quality and can help the team in more different ways.

An important distinction arises between SR and AAS players. Even if SR players are usually versatile and can affect the game on both side of the floor, they are tend to have a weak passing game, which brings to lowers AST numbers. Moreover, they are usually good, but not as good as AAS players in PPG.

While the number of the positions is the same as the classical positions (i.e., 5), the new ones group the players differently. Players who were categorized within the same classic position are now mapped into different clusters. Moreover, every cluster includes players from different old classic positions. Table 6 reports the 5 identified new positions, with examples of players belonging to each category.

As emerged from Figure 6, the clusters of the SOM output layer are not completely separated. Therefore, it is possible for a player to be mapped in a position that is halfway between two groups. In this case, the relative distance between mapped players on the output layer can be used to compare two elements; by analyzing their relative positions it is possible to better characterize the meaning of the defined positions.

Table 6: The 5 new positions derived from the analysis, with examples of players belonging to each category.

Position	Short	Players
All-Around All Stars	AAS	LeBron James (LBJ23), Kevind Durant (KD35), Paul George (PG13), Stephen Curry (SC30), James Harden (JH13)
Scoring Backcourt	SB	Kobe Bryant (KB24), Damian Lillard (DL0), Kyrie Irving (KI2), Dwyane Wade (DW3)
Scoring Rebounder	SR	Marc Gasol (MG33), Carmelo Anthony (CA7), Karl-Anthony Towns (KAT32), Anthony Davis (AD23), Blake Griffin (BG32)
Paint Protector	PP	DeAndre Jordan (DJ6), Andrew Bogut (AB12), Steven Adams (SA12), Kenneth Faried (KF35)
Role Players	RP	Marco Belinelli (MB3), J.J. Redick (JJ4), Harrison Barnes (HB40), Jabari Parker (JP12), Avery Bradley (AB0)

5 Discussion

5.1 Comparison of two players

The relative distance between players on the map shown in Figure 6 allows to assess the result of the clustering and, in a practical usage, to compare the players. As an example, the position of two players will be analyzed: DeAndre Jordan (DJ6), who is assigned to the Paint Protector position (PP), and Damian Lillard (DL0), mapped in the Scoring Backcourt position (SB).

Figure 7 shows the positions of the two players on the map. The PP category is in the lower part of the MDS two-dimensional map, while SB is in the upper part. The two players are placed in opposite positions of the output layer, on the upper and lower edges of the central area of the map. This distance captures the great difference between their statistical profiles, which is due to their different ways to play the game. This is also confirmed by very low mutual membership coefficients (the extent to which each player can be considered to belong to the cluster of the other) assigned by the fuzzy clustering procedure: 0.11 and 0.18 for DeAndre Jordan and Damian Lillard, respectively. Although having a moderate membership to cluster RP (0.23), Lillard is undoubtedly assigned to the SB category, with a membership coefficient of 0.58. In fact, he is a player who can significantly score points and generate many assists. He plays far from the basket and has low values of TRB and BLK. On the other side, Jordan is assigned, with a high membership coefficient (0.72), to the PP position, which characterizes players who usually have low values in statistical parameters as PPG and AST, but are very good at rebounding and blocking shots.

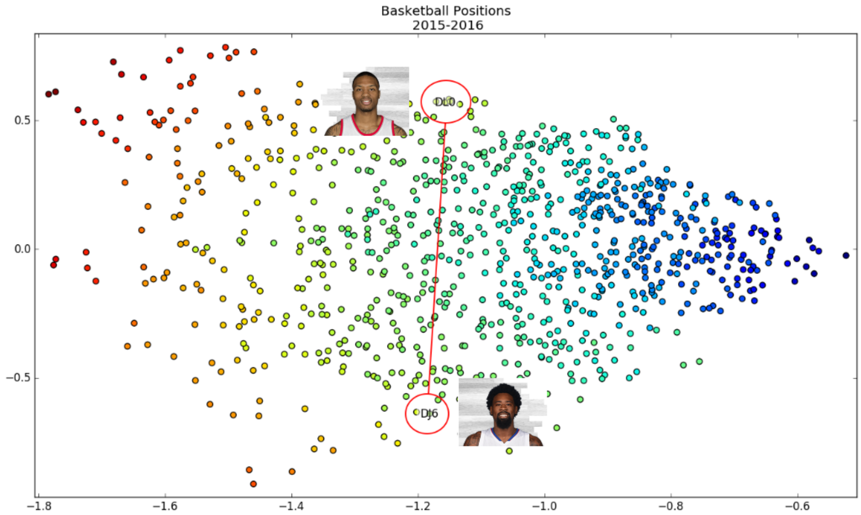

Figure 7: Output layer of the SOM used to compare the classification of Damian Lillard and DeAndre Jordan (photos from NBA.com/stats (2016b)).

To better understand the difference between their statistical profiles, Figure 8 reports the values related to their statistics during the 2015-2016 regular season. The graphical comparison shows how Lillard has almost the double of Jordan's PPG, and averaged 6.8 AST w.r.t. Jordan's 1.2. On the other hand, Jordan's TRB and BLK are more than 3 times higher w.r.t. Lillard ones.

This comparison highlights that Lillard and Jordan play basketball in two different ways. Moreover, it is confirmed that the relative distance between the players in the map is effective in capturing the difference between their statistical profiles.

5.2 Two examples to explore the input space between clusters

Once we verified that the relative distance between players represents a good metric for showing the difference between the way they play, it is worth to discuss the scenario where a player might be mapped half-way between two clusters This situation is due to the output of the SOM algorithm, which may not provide very well-separated clusters, as confirmed also by the fuzzy clustering algorithm. Assuming that the distance between player shows the difference between their statistical profiles (e.g., as in the case of Jordan and Lillard who are mapped very far from each other), there should be some players mapped in between, who have an intermediate way to play basketball.

To show how to analyze this situation, the same example of Lillard and Jordan is used. The area between the two players in the MDS two-dimensional representation is considered, where two more players have been mapped. The statistical profiles evolve while moving through the map, from bottom to the top, gradually changing from one extreme to the other. The two considered players are Karl-Anthony Towns (KAT32) and Carmelo Anthony (CA7), the former assigned to cluster 1 with a still high membership

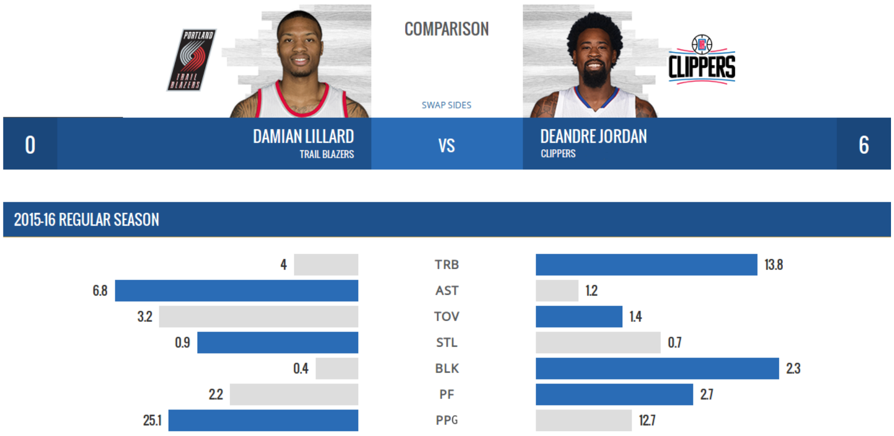

Figure 8: Comparison between the statistics of Lillard and Jordan, Source NBA.com/stats (2016b).

Table 7: From Paint Protector to Scoring Backcourt Statistics

Player	TRB	BLK	AST	STL	TOV	PF	PPG
DeAndre Jordan	13.8	2.3	1.2	0.7	1.4	2.7	12.7
Karl-Anthony Towns	10.4	1.7	2.0	0.7	2.2	3.0	18.3
Carmelo Anthony	7.7	0.5	4.2	0.9	2.4	2.5	21.8
Damian Lillard	4.0	0.4	6.8	0.9	3.2	2.2	25.1

coefficient (0.77), the latter with a genuinely hybrid profile among the three clusters (membership coefficients of 0.39, 0.27, 0.34, respectively) which confirms his positioning "in the middle" among the other considered players.

Figure 9 shows the positions of the 4 players on the MDS two-dimensional map.

An analysis of the statistical profiles was necessary to investigate the evolution of their statistics across the map. The statistical profiles of the 4 players are reported in Table 7 and graphically represented by means of radial plots (Figure 10) and bar charts (Figure 11).

The radial plots clearly show the different profiles of the 4 players. Both the graphs show the change of the statistical profiles of players, which move from the bottom part of the map towards the upper part: PPG and AST (represented in the bar chart with blue and orange bars, respectively, and in the first two rays of the radial plot) gradually increase while moving from the player mapped in the bottom part of the map, i.e., Jordan, to the player mapped in the upper part of the map, i.e., Lillard. On the other hand, the values of TRB and BLK (yellow and green bars in the bar chart, and third and fourth ray of the radial plot, respectively) decrease along the same path. Towns, who is mapped close to Jordan, has better values in PPG and AST than Jordan, but worse than Lillard and Anthony, while he has worse values of TRB and BLK than Jordan,

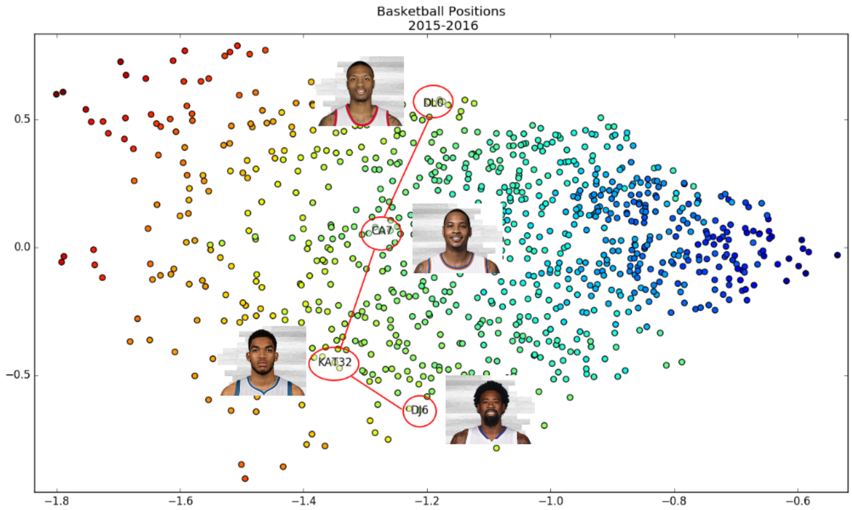

Figure 9: Players from Paint Protector to Scoring Backcourt positions, images of players are taken from Source NBA.com/stats (2016a)

but better values with respect to Lillard and Anthony. Similarly, Anthony is mapped close to Lillard; he is a worse scorer and passer than Lillard, but better than Towns and Jordan, while he is a better rebounder and shot blocker than Lillard, but worse than Towns and Jordan. The way the other variables (STL, TOV, PF) change when moving from one player to the other is more puzzling, as shown by the radial plots, and can be commented as follows. STL has a non-gradual increment, as it presents very low values for the first two players (Jordan and Towns) and suddenly moves to very high values for the other two (Anthony and Lillard). Also TOV has a non-gradual increment, but with different features: it strongly increases when moving from Jordan to Towns, then it remains quite constant for Anthony and it strongly increases again when moving to Lilllard. Finally, PF firstly increases, when moving from Jordan to Towns, then it progressively decreases. This patterns reflect the complexity of the seven-dimensional input space.

6 Concluding remarks

The experiments and the analysis carried out on the relative distances between players prove that the mapping algorithm is an effective descriptor of the difference between players statistical profiles. This also corroborates the validity of the new positions, which are able to describe the way a player plays basketball in a more effective and objective way w.r.t. the classic basketball positions.

Despite the good results, in the new classification there are some differences between players of the same cluster. However, this is something that can not be completely avoided, since each player will always have slightly different statistics w.r.t. the others.

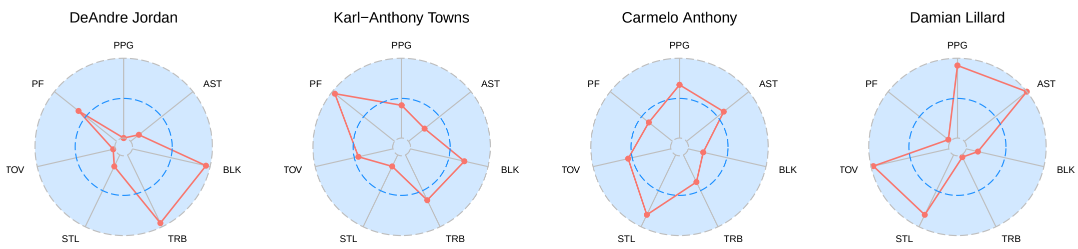

Figure 10: Radial plots of the statistics PPG, AST, BLK, TRB, STL, TOV and PF of the 4 considered players; data of Table 7, standardized (average values denoted by the blue dashed lines)

As a conclusion, the new positions have been proved to be able to describe the main features in the statistical profiles of players assigned to a certain position, which was the expected result of the processing based on the SOM clustering.

6.1 Limits and future developments

The method developed in this paper showed interesting results that could be used as a starting point for further analysis and development of this work. The study has been carried out considering NBA players only and using 7 statistical categories to describe their profiles.

On one hand, the application of the proposed method could be extended to non-NBA players, i.e., by applying the method to any other league in the world. This future work would allow to possibly verify the approach to define new positions in basketball at any level.

On the other hand, a future improvement of this study could also increase the number of statistical variables considered to describe players profiles, adding new and more sophisticated data that nowadays are relatively easy to retrieve. This idea should consider a trade-off, since the more sophisticated the data, the more precise the analysis. However, some data may be available for a limited number of players or they may not be easy to obtain. This is the case, for instance, when the same method is applied to different leagues, where the attention to statistical figures may be much lower w.r.t. NBA. In other words, using sophisticated parameters may not be applicable to the generic league. In addition, when several variables are considered, a preliminary variable selection or a variable clustering may be needed (see for example Kuentz et al., 2015; Vigneau, 2016).

Finally, while the current analysis considers the player "as a whole", another approach could perform the analysis by taking into account the actual positions of the player, i.e., by dividing the behavior of different players to assess his defensive and offensive impact. A coach would thus be able to organize the lineup that fits the need of any specific situation during the game, or through the games during the season.

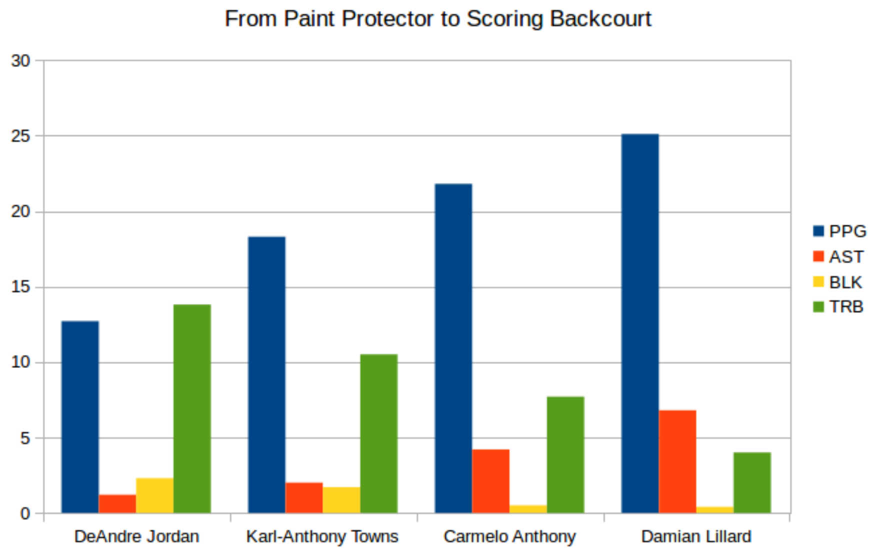

Figure 11: Bar chart of the statistics AST, BLK, TRB and PPG of the 4 considered players, data reported in Table 7

Acknowledgement

Research carried out in collaboration with the Big&Open Data Innovation Laboratory (BODaI-Lab), University of Brescia (project nr. 03-2016, title ”Big Data Analytics in Sports”, `www.bodai.unibs.it/BDSports/`), granted by Fondazione Cariplo and Regione Lombardia.

References

Alagappan, M. (2012). From 5 to 13: Redefining the positions in basketball. `http://www.sloansportsconference.com/content/the-13-nba-positions-using-topology-to-identify-the-different-types-of-players/`. Last access: 20-08-2016.

Bezdek, J. C. (1973). Cluster validity with fuzzy sets. *Journal of Cybernetics*, 3(3):58–73.

Bezdek, J. C. (2013). *Pattern recognition with fuzzy objective function algorithms*. Springer Science & Business Media.

Bianchi, F. (2016). *Towards a new meaning of modern basketball players positions*. Master Degree Thesis in Computer Science and Multimedia, University of Pavia, Italy.

Butts, D. (1986). NCAA adds three-point basket. *The Bryan Times*.

Campello, R. J. and Hruschka, E. R. (2006). A fuzzy extension of the silhouette width criterion for cluster analysis. *Fuzzy Sets and Systems*, 157(21):2858–2875.

Carlsson, G. (2009). Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308.

Corsello, A. (2016). The stunning, strange, beautiful game of Manuel Neuer. http://www.espn.com/espn/feature/story/_/page/enterprise-neuer160525/bayern-munich-manuel-neuer-changing-means-goalie. Last access 19-08-2016.

Flannery, P. (2016). Draymond Green is redefining NBA stardom. Even he didn't see that coming. http://www.sbnation.com/2016/2/16/10987022/draymond-green-warriors-nba-unexpected-star. Last access 19-08-2016.

Kaufman, L. and Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis*, volume 344. John Wiley & Sons.

Kohonen, T. (1982). Self-organized formation of topologically correct feature maps. *Biological cybernetics*, 43(1):59–69.

Kuentz, V., Lyser, S., Candau, J., and Deuffic, P. (2015). Clustofvar-based approach for unsupervised learning: Reading of synthetic variables with sociological data. *Electronic Journal of Applied Statistical Analysis*, 8(2):170–197.

Lynch, J. (1987). High school basketball draws line, adopts 3-point rule. *Los Angeles Times*.

Mahmood, Z. (2015). How 'The Makelele Role' redefined English football. http://www.sportskeeda.com/football/how-the-makelele-role-redefined-english-football. Last accessed: 19-08-2016.

Monson, S. (2012). Mobile Quarterbacks Redefining the Position. http://bleacherreport.com/articles/1251778-mobile-quarterbacks-redefining-the-position. Last access: 19-08-2016.

NBA.com/stats (2016a). NBA.com/stats. http://stats.nba.com/, Last Access 2016.09.02.

NBA.com/stats (2016b). On/Off court comparison. http://stats.nba.com/vs/#!/?PlayerID=203081&VsPlayerID=201599, Last Access 2016.09.02.

Sanders, S. (1981). 22 will get you 3. *Spartanburg Herald*.

Vigneau, E. (2016). Dimensionality reduction by clustering of variables while setting aside atypical variables. *Electronic Journal of Applied Statistical Analysis*, 9(1):134–153.

Winkler, R., Klawonn, F., and Kruse, R. (2011). Fuzzy clustering with polynomial fuzzifier function in connection with m-estimators. *Appl. Comput. Math*, 10(1):146–163.