

Research on Basketball Defensive Strategy Modeling Based on Movement Trajectory Data

Hongze Gao^{*}

RCF Experimental School, Chaoyang District, Beijing, 100029

^{*}Corresponding Author, Email: gaohongze@rdfzcygj.cn

Abstract

In the process of half-court defense, the defending team often has a weak link in the defense due to the slow rotation speed, resulting in a certain gap. In order to formulate a reasonable defense plan, this paper uses the NBA tracking data and builds a defense scheduling model based on MIP. In the study, the planar coordinates of the players are extracted, and the movement speed is analyzed. Voronoi diagram is used to calculate the open space of the offensive players, and the risk of defensive breakdown is calculated. In order to achieve the goal of minimizing the total rotation time of the whole team, Python is used to calculate and solve the problem. The results show that the defense plan formulated by the model can make up for the shortcomings of the traditional man-to-man defense. When facing a ball handler with a higher threat, the system can automatically generate a double team strategy, and the offensive players in the low-efficiency scoring area can be left open. By introducing physical coordinates and spatial geometric data, this paper can provide reference for the formulation of defense plan.

Key words

Basketball Analytics; Mixed-Integer Programming; Defensive Strategy; Voronoi Diagram

1. Introduction

1.1 Research Background

In recent years, optical tracking technology has been widely used, and the recording methods of basketball

game data have changed^[1]. The coordinates of the players (x, y) can be obtained in seconds, and the movements of the players can be analyzed by mathematical methods.

In the game of basketball, defense is the most critical content to test the cooperation ability of the whole

team. In the process of attack, the defense should be able to make reasonable decisions in a short time, which requires "who guards whom". In the process of daily competition, the coach should make a reasonable arrangement before the game, and the players should also have rich experience. However, the speed of the player's movement and the change of the space on the court are unpredictable, and the rigid tactical presetting cannot meet the requirements of the game, and the defender cannot arrive at the specified position^[2].

In the process of organizing and analyzing the trajectory data of basketball games, how to use mathematical models to reasonably allocate defensive resources without the guidance of tactical arrangements, and to calculate the space occupied by the offensive side, the distance between players, the speed of movement and other factors, it is necessary to solve the problem of operations research.

1.2 Current Research Status

At present, the academic circles and the industry have made efforts to explore the evaluation of basketball defense data, but there are still some shortcomings.

In the initial stage of spatial analysis, macro indicators were usually used to describe the geometric characteristics, and the convex hull algorithm was introduced into the analysis of basketball, which can be used to calculate the area of the polygon formed by the players on the court, and to analyze the defensive space and the tightness of the team formation^[3]. However, the convex hull area is a macro indicator, which cannot accurately grasp the space occupied by the offensive player, and cannot guide the defensive personnel to make reasonable arrangements.

At present, the EPV model has attracted the attention of more and more people, which is a model with strong cutting-edge nature. The model uses Markov chain to calculate the shooting and passing expectations, and divides the court into several blocks^[4]. In the study, some scholars believe that the calculation of the vertical projection distance between the passing route and the defender can get the "passing lane congestion". Although these models have achieved certain results in the evaluation, the probability model is more complex, and it is difficult to achieve the goal of deterministic defense.

In the process of optimizing the defense strategy,

the preset tactics are usually adopted, including "double-teaming" and "switching", and the physical movement cost is measured according to the set tactics. This method can play an important guiding role in evaluating the effectiveness of specific tactics. However, the basketball defense is inevitably affected by the position of players and the physical space, and the defense strategy will change dynamically. Therefore, this paper does not use the method of introducing preset tactics, but explores from the physical and geometric level, and regards the defense strategy as an optimal solution, which can minimize the positioning time and maximize the offensive pressure.

1.3 Research Objectives

In this study, the MIP method is used to analyze the real tracking data of NBA, and the optimal assignment model is constructed to make the basketball defense more dynamic. The model not only considers the running speed of the player, but also quantifies the offensive space, and the purpose of this research is to minimize the total cost of the system, minimize the defensive rotation time of the whole team, and formulate a reasonable plan to guide the defensive rotation of the whole team.

1.4 Innovations of This Study

The innovations of this study are reflected at three progressive levels: data representation, optimization modeling, and tactical response.

From static defensive evaluation to trajectory-based spatial representation. Instead of relying only on macro-level indicators or preset tactical labels, the study directly uses NBA optical tracking coordinates and instantaneous movement information to represent the locations and movement states of players. Voronoi partitioning is further introduced to quantify the control area of each offensive player, allowing defensive pressure to be evaluated at the individual spatial level rather than only at the team-formation level.

Integration of physical movement cost and spatial threat within a unified MIP framework. The model simultaneously considers the time required for defenders to complete rotations and the penalty associated with leaving threatening offensive space insufficiently covered. By combining the defensive rotation-time matrix

with spatial threat weights and defender-assignment constraints, the problem of ‘who guards whom’ is transformed into a solvable resource-allocation problem that balances physical feasibility and defensive risk.

Dynamic generation of defensive adjustments without predefining specific tactics. Rather than hard-coding switching or double-teaming as fixed tactical rules, the model allows these behaviors to emerge from the optimization solution as player positions and spatial threats change. The sensitivity analysis of an unguarded player's cut toward the basket further shows how the model can trigger tactical restructuring when the spatial risk crosses a critical level, thereby improving the adaptability of defensive decision-making.

2. Research Methodology Design

2.1 Introduction to Methodology

(1) Data collection and processing: Python is used to parse the real data set, extract the two-dimensional coordinates of the offensive and defensive players, find out the position of the ball handler, and use the backward difference method to calculate the initial movement speed of the defensive players, so as to obtain the basic parameters required for the physical model.

(2) Construction of mathematical model: MIP method is used to make scientific decisions in the process of matching and assigning defensive players, and the objective function is set to the lowest possible cost, which requires the shortest possible defensive running time and the lowest defensive breakdown penalty. In the process of setting constraints, the Voronoi diagram algorithm is used to calculate the risk of breakdown, and the uniform acceleration equation is used to calculate the running time.

(3) Model solving and verification: PuLP library is used to solve the model, and the optimal solution can be obtained. According to the model, the shortest team defense rotation time and the lowest breakdown penalty can be obtained. In order to enhance the stability of the model, the time cost weight and spatial penalty weight are adjusted, and the comparative experiment is carried out.

(4) Result analysis and application suggestions:

In this study, the internal logic of the optimal defensive plan is analyzed, the significance of the plan is analyzed, and reasonable suggestions are put forward for the formulation of the defensive rotation strategy of the team.

2.2 Problem Description

In the process of half court set defense, the defending team needs to solve a complex dynamic resource allocation problem. There are 5 offensive players and 5 defensive players on the court. The offensive players are constantly moving and passing the ball, and the defensive players need to formulate a reasonable defense plan in a short time.

Defenders should not just guard the nearest opponent, because this will lead to the formation of unguarded areas, and the shooting threat of the other side will increase; if they run a long distance to defend, the running time will be relatively long, and the offensive side can easily complete the shot. In this paper, the problem to be solved is how to make a reasonable match plan, grasp the position of both sides of the players, the speed of the defenders, and the position of the ball holder, and suppress the shooting space of the offensive side.

2.3 Data Sources and Processing

In this study, the data is obtained from the NBA optical tracking dataset. The data set can record the planar coordinates of 10 players and a basketball at a frequency of 25 frames per second.

To transform the massive raw data into core parameters usable by the model, the following data processing steps were performed:

Firstly, extract the coordinate data of a specific game frame, and determine the absolute position of the basketball, so as to determine which half of the court the battle is taking place, and construct the coordinate matrix respectively.

Secondly, the distance between the basketball and all offensive players is calculated according to the Euclidean distance, and the offensive player with the shortest distance to the basketball is identified as the core ball handler, which can play a role in the next step of defense.

Thirdly, the actual instantaneous speeds of the defensive players were extracted. Utilizing the continuous

characteristics of high-frequency data, the backward difference method was adopted to calculate the initial movement speed of each defensive player by dividing the coordinate displacement between the current frame and the previous frame by the single-frame time interval.

2.4 Model Introduction

In this study, in order to solve the problem of basketball strategic planning, the linear programming method is introduced. This is a method in operations research, which can help us find the optimal solution under the premise of setting a series of constraints, and the objective function is linear ^[5]. The mathematical model of linear programming is:

$$\text{Max}Z = C_1x_1 + C_2x_2 + \dots + C_nx_n \text{ s.t.}$$

$$a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \quad a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2$$

$$\dots\dots a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \quad x_1, x_2, \dots, x_n \geq 0$$

In the above formulation, Z denotes the value of the objective function to be maximized; $x_1, x_2, \dots, x_n$ are the decision variables; $C_1, C_2, \dots, C_n$ are the objective-function coefficients that measure the contribution of each decision variable to Z ; a_{ij} denotes the coefficient of decision variable x_j in the i -th constraint; b_i is the right-hand-side constant representing the available resource or required bound of the i -th constraint; n is the number of decision variables and m is the number of constraints. The notation “s.t.” means “subject to,” indicating that the objective function is optimized while all listed constraints are satisfied, and $x_j \geq 0$ specifies the non-negativity requirement for the decision variables. In the basketball-defense problem developed later, these abstract decision variables are instantiated as defender-to-offensive-player assignments, while the coefficients and constraints are constructed from player distances, movement time, spatial-threat measures, and coverage requirements.

In the process of constructing the spatial defense risk model, this paper chooses Voronoi diagram as a tool, which is a spatial geometric segmentation tool ^[6]. Through this tool, the continuous space can be divided into multiple control areas, and the absolute position of

the player on the court is the basis for division. From the mathematical point of view, Voronoi polygon can be used to describe the offensive player, and all points in this polygon are closer to the offensive player than to other players on the court. This is a geometric segmentation, which can accurately describe the control area of each offensive player.

Linear programming is a method used to solve problems in a variety of fields, such as transportation, production planning, resource allocation and so on. Generally, the simplex method is used to solve linear programming problems. If integer variables are introduced into the linear programming problem, it is called mixed integer programming problem, which is solved by cutting plane method and branch and bound method ^[7].

3. Model Solving and Analysis

3.1 Model Assumptions

To simplify the problem, this paper proposes the following basic assumptions:

(1) Two-dimensional motion assumption: It is assumed that the players are in the same plane when they cover the space and run to defend, and the height difference and jumping will not affect the distance.

(2) Assumption of uniform acceleration linear motion: In the process of defense, the defensive players are assumed to move in a straight line, and the acceleration is constant.

(3) No-Obstacle Assumption: All players on the court are regarded as point masses, and they have mass and velocity. When calculating the shortest physical time for the defender to reach the target position, the physical blocking caused by other players is temporarily ignored.

3.2 Model Establishment

3.2.1 Symbol Table

To make the practical meaning of the subsequent model establishment more intuitive, the mathematical symbols used in the model are uniformly defined first:

Table 1. Symbol Table

Symbol	Description
I	Set of defending players
J	Set of offensive players
b	Index of the core ball handler
D_{ij}	Straight-line distance between defender i and offensive player j on the 2D plane
v_{0i}	Initial movement speed of defender i at the current calculation frame
a	Average starting acceleration constant during a defender's close-out
T_{ij}	Defensive rotation time matrix; the time required for defender i to reach offensive player j
A_j	Area of the Voronoi polygon controlled by offensive player j
W_j	Dynamic spatial threat weight of offensive player j
K_j	Number of defenders required to fully restrict the player

3.2.2 Decision Variables

This model contains three types of decision variables, corresponding to defensive assignment, time cost, and spatial risk respectively:

$x_{ij} \in \{0,1\}$: If the system assigns defender i to guard offensive player j , $x_{ij} = 1$; otherwise, $x_{ij} = 0$.

Z_{max} : Represents the maximum running time required for the team to complete the defensive matchup plan.

O_j : Represents the risk of defensive breakdown faced by offensive player j .

3.2.3 Objective Function

The purpose of this model is to formulate a reasonable defense plan for the team and minimize the comprehensive cost. The cost is composed of two parts, one is the cost of defenders running, the other is the cost of defenders positioning, and the cost of not being able to defend against offensive players is also included. The objective function is as follows:

$$\text{Min}Z = \beta Z_{max} + \gamma \sum_{j=1}^5 (A_j \cdot W_j) \cdot O_j$$

The first half of the formula represents the time cost; the system will tend to choose a plan that minimizes the running time of the slowest defender on the team. The second half of the formula is the sum of the defensive breakdown penalties, where $A_j \cdot W_j$ represents the objective threat expectation of the space where offensive player j is located. When there is a risk of leaving this

player open, the system will incur a corresponding penalty. β and γ are weight constants adjusting the proportion of these two costs.

3.2.4 Constraints

(1) Defensive Personnel Allocation Constraint: At any defensive instant, each defender must and can only be assigned to guard one offensive player.

$$\sum_{j=1}^5 x_{ij} = 1$$

(2) Team Defensive Time Constraint: For all effective defensive assignments, the corresponding running time T_{ij} cannot exceed the team's defensive time bottleneck Z_{max} .

$$Z_{max} \geq T_{ij} \cdot x_{ij}$$

(3) Defensive Breakdown Risk Constraint: Used to determine whether the assigned number of defenders is sufficient to offset the offensive player.

$$O_j \geq K_j - \sum_{i=1}^5 x_{ij}$$

3.3 Model Results

In the NBA game, the initial speed and real tracking coordinates of the players are substituted into the model, and the PuLP library is used to solve it. The program can identify the battlefield and accurately grasp the right half court. After solving the mixed integer programming, the spatial division results and the optimal defense plan are obtained, as shown in Figure 1.

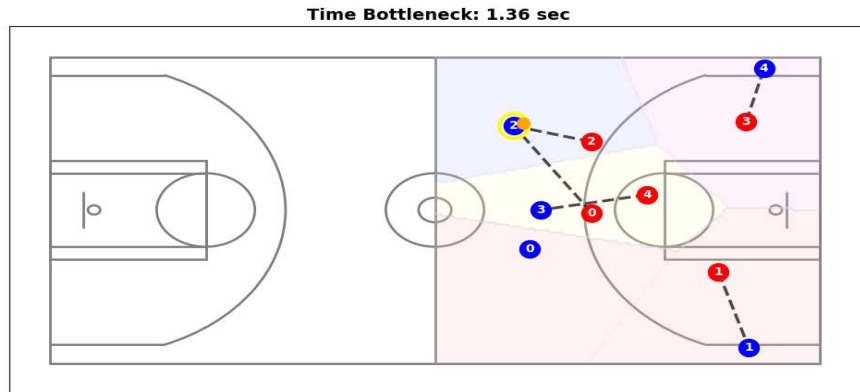


Figure 1. Results obtained from solving

3.4 Results Analysis

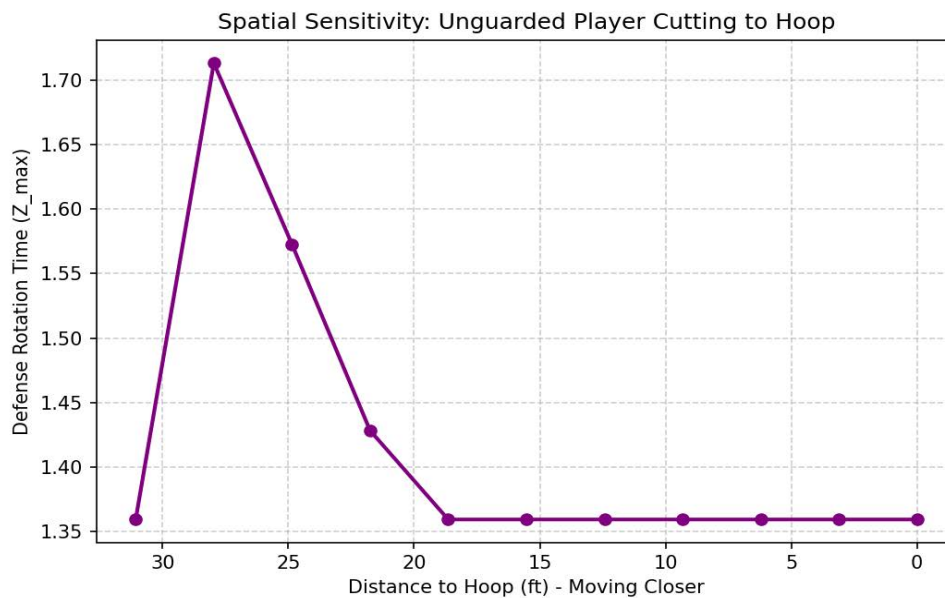


Figure 2. The trend of defensive rotation time varying with the offensive player's cutting distance to the basket

From the above chart, it can be seen that the model can accurately identify the threat to the court, and assign two defenders to the ball handler (Blue No. 2) to double team him. However, due to the limited number of defenders, the system decided to leave Blue No. 0 open, because he was closer to the ball handler. The reason for this decision is that the No. 0 player is in the mid-range area, and the scoring rate is relatively low, which is lower than that of No. 4 and No. 1. It can be seen that the model can allocate defensive resources rationally

according to the scoring efficiency of space.

In order to ensure the effectiveness of the model, this study focuses on simulating the situation that the player without defense can cut to the basket quickly, and records the dynamic feedback of the system. It can be seen from Figure 2 that the defensive time will change with the change of cutting distance. When the player is in a low-efficiency area, the system will not interfere with him, and the defensive time will remain unchanged at 1.36 seconds. When the player cuts into a high-risk

area, the offensive threat will increase sharply, and the model will be forced to organize other defenders to provide help defense. In this case, the formation will be stretched, and the team's defensive running time will increase sharply to 1.70 seconds. With the deepening of cutting, the distance between the two sides becomes closer, and the maximum running time continues to decline.

In order to grasp the evolution of tactical restruc-

turing, the dual-axis chart in the third chart is analyzed. The blue dashed line in the chart shows that when the cutting distance reaches the critical point, the player's defensive breakdown risk drops rapidly from 1 to 0, which means that the model has forced the defensive switch. The red line is a solid line, which means that the team has to pay the highest defensive cost in the process of tactical transformation.

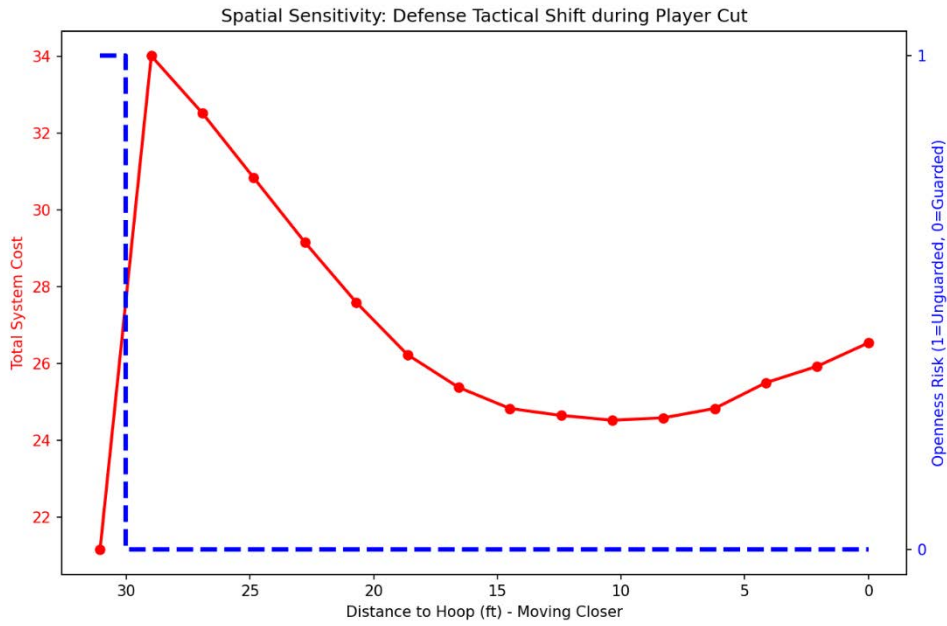


Figure 3. The evolution of defensive breakdown risk and total defensive cost during tactical restructuring

4. Conclusion and Outlook

In this research, the operation research theory is introduced into the basketball set defense, and the mixed integer programming model is constructed to achieve the purpose of application. The difference between this model and traditional static model is that the latter depend too much on preset tactics, while this model uses the NBA optical tracking data to analyze and obtain the real data. In this case, the model introduces the time consumption matrix into it, and the movement is accelerated in a straight line, and the constraints are set according to the Voronoi diagram and the scoring efficiency of each area. In this study, the spatial sensitivity test is carried out according to the physical coordinates of the offensive players, and the adaptability of the system

is analyzed.

Using PuLP library in Python to solve the model, this paper formulates the optimal matching plan and minimizes the comprehensive defensive cost. It is not difficult to see that this plan can not only control the running time of the team's defense, but also make reasonable tactical arrangements according to the actual situation of the game, such as sending two people to defend the ball handler and leaving the mid-range player in the open space, which fully reflects the significance of mathematical programming in the analysis of complex sports competition. This research result is enough to attract the attention of coaches, which can help coaches make scientific decisions in the game.

It should be noted that the model constructed in this paper has a certain limitation. The current research

conditions and the linear solution framework are relatively limited, and only the exponential decay of spatial distance is used to calculate the defense in the process of processing the passing channel, but it does not reflect the characteristics of the passing network, while actual defense forms a complex network through body positioning and jumping. In the current model, it is assumed that the players' acceleration is a constant, but it cannot reflect the physical consumption pattern of the players in the whole game. At present, the model has certain limitations, which requires the use of graph theory algorithm to simulate the geometric occlusion, and to introduce nonlinear fatigue function. At the same time, the macro-tactical system of the team should be taken as the boundary constraint, and the Markov chain method should be introduced to describe the dynamic transition of the offensive ball flow and the movement state, so as to make the defensive configuration model more consistent with the real game.

References

- [1] P. Lucey, A. Bialkowski, P. Carr, Y. Yue, and I. Matthews, "Analyzing team movement in basketball using tracking data," in *Proceedings of the 8th MIT Sloan Sports Analytics Conference*, vol. 2, 2014.
- [2] L. Lamas, F. Santana, M. Heiner, C. Ugrinowitsch, and G. Fellingham, "Modeling the offensive-defensive interaction and resulting outcomes in basketball," *PLoS One*, vol. 10, no. 12, p. e0144435, 2015.
- [3] A. Franks, A. Miller, L. Bornn, and K. Goldsberry, "Characterizing the spatial structure of defensive skill in professional basketball," *The Annals of Applied Statistics*, vol. 9, no. 1, pp. 94-121, 2015.
- [4] D. Cervone, A. D'Amour, L. Bornn, and K. Goldsberry, "A multiresolution stochastic process model for predicting basketball possession outcomes," *Journal of the American Statistical Association*, vol. 111, no. 514, pp. 585-599, 2016.
- [5] Y. Hu and Y. Guo, *Operations Research Tutorial*, 4th ed. Beijing, China: Tsinghua University Press, 2011.
- [6] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. Chichester, UK: John Wiley & Sons, 2000.
- [7] R. E. Burkard, M. Dell'Amico, and S. Martello, *Assignment Problems*. Philadelphia, PA: Society for Industrial and Applied Mathematics (SIAM), 2012.