Visual Analytics of Ball Handlers' Decisions in Basketball Games

Yihong Wu* [1] Ziao Liu† [2] Liqi Cheng‡ [1] Moqi He§ [1] Dazhen Deng¶ [3] Xiao Xie‖ [2]
Hui Zhang** [2] Yingcai Wu†† [1]

[1] State Key Lab of CAD&CG, Zhejiang University
[2] Department of Sports Science, Zhejiang University
[3] School of Software Technology, Zhejiang University

ABSTRACT

In basketball, decision-making is one of the core skills for players. For example, when a player is holding the ball, the success of the team's offense is primarily determined by her/his decisions (i.e., pass, shoot, or dribble) in response to the dynamics of the game. Understanding players' decision-making processes in changing game situations can help coaches develop effective strategies, which is critical for the success of a team. However, the decision-making process is influenced by various factors (e.g., player's playing style, opponents' defense, and time remaining), making understanding a challenging problem. In this study, we propose HoopScouter, a visual analytics system to help understand ball handlers' decisions in basketball games. Based on a careful investigation of the analysis requirements, we first introduce a representation learning method that characterizes ball handlers' decision-making styles. We then design a sketch panel with integrated time information to support exploration of player decisions under similar game scenarios. Facet views and coordinated interactions are also provided to identify the strengths and weaknesses of the ball handler's decision-making, and to understand when and why ball handlers would make certain decisions. To validate the effectiveness of HoopScouter, we conduct two case studies on real-world basketball games and receive positive feedback from domain experts.

Index Terms: Human-centered computing—Visualization—Visualization application domains—Visual analytics

1 INTRODUCTION

Decision-making is a fundamental skill for basketball players. Players with the ball in hand always have to make decisions about whether to shoot directly, pass to teammates, or continue dribbling, which can have a great impact on the game result. Although important, making a wise decision is difficult since the player has to fully consider the dynamic scenario on the court, such as the positions of teammates and opponents, the score, and the time remaining. This complex decision-making process, however, cannot be directly reflected in the result-oriented match statistics, such as points, assists, and turnovers. Hence, to help players improve their decision-making skills and make better decisions, coaches and analysts need to browse lots of videos to identify exemplars, which is cumbersome and labor-intensive.

With the broad deployment of optical tracking systems [1] in recent years, fine-grained tracking data is promising for addressing the inefficiency of traditional analysis. Based on the tracking data, various machine learning models can be applied in decision prediction [18, 24] and evaluation [10, 54] to help coaches instruct the ball handler to make appropriate decisions. Despite its usefulness, coaches may encounter two problems when using the models to analyze decisions. First, the prediction results are difficult to verify by actual games. This causes coaches to rely heavily on their own experience to determine whether the predictions are reliable, thus decreasing their trust in the model. Second, the model outputs are usually not integrated directly into the specific context of a basketball game, making it difficult to give coaches an intuitive overview of the decisions (e.g., which decision is better at crucial moments of the game).

To tackle these problems, we first extend the traditional analysis framework with the tracking data and a sketch-based search method. The search method imitates the way coaches draw a basketball play on the whiteboard, allowing coaches to visually depict a game scenario and find the historical decisions they want to analyze. Then we propose a visual analytics approach that presents the searched decisions through a court-based visualization and compares the decisions from multiple perspectives. The approach can help identify suitable decisions under a specific situation and explore the decision diversity between players or across game contexts. We face three main challenges when establishing such a decision analysis framework.

Quantitative characterization of the playing styles. Players' playing styles can roughly reflect their decision-making characteristics (e.g., pass-first and score-first players). If players with similar playing styles are grouped together, it can guide experts to quickly find the significant differences in decisions. However, the current definition of playing styles is primarily derived from coaches' subjective opinions based on long-term observations. As a quantitative characterization of playing styles is demanded, Decroos et al. [14] proposed a method to describe soccer players' playing styles via on-ball movements. This method cannot be simply applied since it discards important context information for basketball, such as how the player copes with the fast-changing defense, which is the challenge of the characterization.

Visual comparison of multivariate decisions. Ball handlers' decisions contain various types of multivariate information, including categorical data (e.g., decision type), numerical values (e.g., scoring rate), and spatial distribution (e.g., target location of a pass). Coaches need a comprehensive overview to summarize the general pattern of decisions, and a visual comparison to understand the diversity between players or in different game contexts. Yet various attributes and multiple spatial areas to be compared largely hinder coaches' perception, leading to the requirement of tailored decision visualizations.

Interactive system for analyzing decisions. An interactive system is required to provide a flexible exploration of historical decisions of interest. However, this can be challenging for two reasons. First, searching for decisions based on coaches' sketches is difficult due to the large number of games and the complex structure of player movements. Second, few existing sports visualizations involve the analysis of players' decisions. How to design a system to help analysts understand players' decisions and inspect various influences on the decisions remains an open question.

To address the first challenge, we propose an approach based on Word2Vec [33, 34] to capture the degree of decision-making similarity across players. By representing players as vectors, this approach offers a quantitative measure of the playing styles. To address the

*e-mail: wuyihong@zju.edu.cn
†e-mail: liuziao@zju.edu.cn
‡e-mail: lycheecheng@zju.edu.cn
§e-mail: moqiHe@zju.edu.cn
¶e-mail: dengdazhen@zju.edu.cn
‖e-mail: xxie@zju.edu.cn
**e-mail: zhang_hui@zju.edu.cn
††e-mail: ycwu@zju.edu.cn
‡‡Xiao Xie is the corresponding author.

second challenge, we design a series of coordinated visualizations and combine them with a basketball court to visualize and compare the statistical and spatial information of ball handlers' decisions. Based on the design, we further implement HoopScouter, an interactive system that facilitates the visual understanding of ball handler's decisions. The main contributions of this work are as follows:

◇ Proposing a representation learning method to characterize the playing style of basketball players.
◇ Implementing a comprehensive system that supports interactive sketch-based retrieval and detailed analysis of player decisions.
◇ Demonstrating the effectiveness of HoopScouter through two case studies based on the tracking data of over 600 basketball games.

2 Related Work

In this section, we present relevant studies, including basketball analysis and visualization, sketch-based search in sports analytics, playing style characterization, and visual analytics for decision-making.

2.1 Basketball Analysis and Visualization

Fine-grained tracking data offers the potential for in-depth analysis of basketball [1]. Various methods use tracking data to solve challenging problems, such as tactic recognition [11, 59, 62], performance assessment [20, 31, 54], and game simulation [12, 25, 50]. Closely related to our work, POINTWISE [10] incorporated tracking data into a Markov model to predict and assess player decisions. However, integrating game context or presenting different players' decision-making tendencies is not the focus of their study, which may pose difficulties in understanding these processes. To address this, our work employs tailored visualization to reveal players' decision patterns and uncover the effects and reasons behind the decisions of ball handlers.

As visualization can be more accessible than statistical models, it has become an integral part of sports analysis [17, 40], including table tennis [61, 65, 67], tennis [15, 41, 42], and soccer [3, 39, 46, 56]. As for basketball, we classify existing studies into two categories: visual game storytelling and analytics. Visual game storytelling aims to promote viewers' understanding of the game process [13, 30, 55]. For example, Zhi et al. [69] proposed two exploratory prototypes that present key game events from the perspectives of basketball fans and sportswriters. Lin et al. [27] designed innovative visualizations to embed extra stats into live videos, improving the audience's viewing experience. Visual game analytics helps evaluate player performance and team tactics [6]. For example, Goldsberry et al. [21] proposed a court-based visualization to quantify players' shooting performance in games. Losada et al. [29] developed a visual analytics system that reveals both individual and team performance through multiple match statistics. Wu et al. [66] combined an interpretable model with visualization techniques to assess and explain the value of off-ball player movements. While these methods involve player decision-making to varying degrees, how players make trade-offs between different decisions in real game scenarios has not been explored. This study aims to fill that gap with our interactive system, HoopScouter.

2.2 Sketch-based Search in Sports Analytics

Sketching on the whiteboard is an intuitive way for coaches to convey complex play actions [43]. Thus, many studies take sketches as the input to search for target situations. For example, Legg et al. [26] contributed a visualization system, in which analysts can sketch a path showing the team's movement and retrieve relevant video clips in rugby games. Given the importance of game context, Stein et al. [57] integrated trajectory shape, player roles, and contextual information into the search process. However, the search input for both methods is a single trajectory, which is insufficient for basketball plays as they typically involve multiple players. To search by multiple trajectories, Shao et al. [53] adopted a feature-based similarity function and reduced the search space through an event-specific filtering process. Sha et al. proposed a tailored method that aligns multi-agent data [52] and supports the retrieval of basketball plays at interactive speed [51]. Seebacher et al. [49] designed a virtual magnetic tactic board to retrieve similar team behaviors from a huge amount of football match tracking data. Despite the usefulness, these multi-trajectory search methods can only compare the trajectories within time windows of the same length, limiting the number of matched samples. To address this, we provide an approach based on spatial indexing to make the search possible without strict temporal alignment.

2.3 Playing Style Characterization

A quantitative description of playing style can reflect a player's in-game decision-making tendency, which is valuable for player scouting and tactical planning. Many researchers have studied this challenging task. For instance, Bialkowski et al. [7] investigated the soccer team's usage of different formations and captured the playing style from it. For individual players, Gyarmati et al. [23] quantified the uniqueness of soccer players by producing their in-game movement vectors from an event-based dataset. In basketball, Franks et al. [19] aggregated players' shooting locations to acquire the shooting habits of different players. Despite their usefulness, these studies are based on only one type of decision, which is not a comprehensive characterization of playing styles. To address this, Decroos et al. [14] constructed a vector for each type of on-ball decision (e.g., passes, shots, dribbles) and combined these vectors to quantitatively describe the characteristics of soccer players. Nevertheless, the approach still neglects the game context (e.g., the defense and score difference) in which players make decisions. Hence, we propose a representation learning method based on Word2Vec [33, 34] that integrates contextual information into the characterization of basketball players' decision-making. The representations learned can help experts quickly find the players who tend to make similar decisions in similar game contexts.

2.4 Visual Analysis of Decision-Making Processes

The analysis of decision-making processes has become one of the most popular research topics in the visualization community. Most studies focus on multiple-criteria decision-making (MCDM), assisting users in seeking the best solution under multiple constraints. Available decision support tools, like ValueCharts [9] and ValueCharts Plus [5], adopt a table-based stacked bar chart to rank solutions by weighted summary scores. Later, LineUp [22] extends the tools with slope graphs and allows interactive refinement of attribute weights. To assess the sensitivity of a solution to weight variations, WeightLifter [38] offers a comprehensive exploration of the weight space. For non-expert users, Podium [60] allows them to rank alternative solutions manually, then automatically calculates appropriate weights using machine learning methods. In addition, a variety of visualization systems are designed for domain-specific scenarios, such as urban computing [4, 28, 63] and epidemic control [2, 68]. Different from the studies above, our system focuses on uncovering players' decision-making processes at specific moments. It provides a visual summary of ball handlers' decisions in a given situation, reflecting the effects of factors that drive their different decision-making processes.

3 Background and System Overview

In this section, we introduce the background of our study, then analyze the domain requirements from expert interviews.

3.1 Background and Concepts

A professional basketball game involves two teams, each with five players on the court. In the NBA, the game usually consists of four periods, and each period is 12 minutes long. At the end of the game, the team that scores more points will win.

- *Ball handlers* denote the players who are in possession of the ball.
- *Possessions* refer to the opportunities for a team to score points. The time limit for each possession is 24 seconds.
- *Shot clock* records the remaining time for a possession.
- *Game clock* records the remaining time for a period.
- *Decisions* refer to the ball handlers' options on offense, including shooting, passing, and dribbling.
- *A situation* refers to a series of player movements within the same time period. In this study, coaches sketch player movement paths to define a situation. Then, situations containing similar movement behaviors will be retrieved from the tracking data. The decisions made at the end of these retrieved situations are our main focus.
- *Playing positions* refer to the players' roles, such as point guards, power forwards, and centers. To distinguish the term from spatial positions, we consistently use *position* to denote the playing position and *location* to denote the spatial position.

- *Game context* consists of external factors that influence the ball handlers' decision-making, such as the defense and shot clock.

3.2 Requirement Analysis

In the past year, we collaborated closely with three domain experts, including one professional basketball coach (E1, who has worked for the women's national basketball team), one senior sports analyst (E2, who has decades of experience in sports data analysis), and his Ph.D. student (E3). Our goal is to identify and compare the decision-making patterns of different players facing various game contexts.

Guided by the nested model for visualization design and validation [35], we first reviewed relevant literature, discussed the domain problems with the experts, and developed a pilot system based on the domain problems. During the downstream problem characterization, we encountered challenges at the domain and abstraction levels. Initially, we considered team-based analysis, but experts noted that a ball handler's decisions are mainly influenced by individual style, making team comparisons potentially inaccurate. Additionally, while experts suggested incorporating player attributes like height, wingspan, and jumping ability into the analysis, these data were missing from the tracking records. Then we iteratively refined the system and user requirements in response to weekly expert interviews. Finally, we compiled six requirements into three aspects as follows.

L. Locating Decisions can help experts search through a large number of games and find the decisions made in a specific situation.

L1 **Depict the critical decision-making situation.** The ball handlers face various decision-making situations in the game. Thus, experts need to accurately depict the situation they want to analyze. The situation generally involves the movement of one or several players—for example, *"when a teammate sets a screen and frees the ball handler."* To depict such a situation, experts can sketch the player locations and their running directions on a whiteboard, which is intuitive and efficient [43].

L2 **Retrieve the decisions made at the end of the situation.** After depicting the desired situation, the experts require an efficient search engine to retrieve concrete samples and acquire the ball handlers' decisions made at the end of these situations. They also need to review the players' movements and decisions in each retrieved sample for verification.

S. Summarizing Decisions can provide experts with an overall understanding of the retrieved decisions.

S1 **Obtain a statistical overview of the retrieved decisions.** As a starting point of the analysis, experts require a statistical overview to understand a large number of the retrieved decisions. The overview should provide key indicators of these decisions: *"What proportion of decisions involve shooting, passing, and dribbling? Which decision is more effective against defenders? What are the outcomes of the different types of decisions?"*

S2 **Examine the spatial distribution of the retrieved decisions.** The target location of a decision, such as where the ball handler passes the ball to, can provide crucial details to evaluate the decisions. For instance, experts can find the locations where the ball handlers frequently pass to, and optimize the decisions based on the corresponding scoring rates. To perform such an analysis, experts should examine how the target locations are distributed over the court and learn the scoring rate at each location.

C. Comparing Decisions can help experts examine the decision-making characteristics of different player groups and identify suitable decisions in different contexts.

C1 **Compare decisions between different groups of players.** Despite facing similar situations, different players may vary considerably in their decision-making. Thus, the experts wish to compare different groups of players from various perspectives and infer the potential reasons for the differences. Players can be grouped by their playing positions (e.g., guards and forwards) or playing styles (e.g., pass-first and score-first players).

C2 **Compare a player group's decisions considering contextual factors.** Contextual factors in games, like the remaining time and score difference, can also affect the ball handlers' decision-making. Hence, the experts want to compare the same player group's decisions in different game contexts and inspect how a contextual factor influences decision tendencies and quality.

3.3 System Overview

Based on these requirements, we designed HoopScouter, a visual analytics system that helps coaches analyze ball handlers' decision-making. The system consists of three components: playing style modeling, situation search, and decision visualization. The playing style modeling component quantifies ball handlers' styles using representation learning, embedding them as vectors in the decision visualization module to highlight players with similar decision tendencies. The situation search engine enables flexible queries based on coaches' sketches, allowing them to retrieve relevant historical scenarios efficiently. Bringing these components together, the decision visualization module provides an interactive platform for customized decision analysis. We implemented the system using PyTorch for modeling, a prefix tree for search, and Vue.js for visualization.

4 Model

In this section, we introduce an efficient model for characterizing the ball handler's playing style with quantitative evaluation.

4.1 Data Description

We use a public dataset from STATS SportVU [1], which contains tracking data and event data for over 600 games during the 2015-16 regular season. In the tracking data, the spatial locations of the ball and players, as well as temporal information such as the game clock, are collected at 25 frames per second. In the event data, each event record is composed of timestamps, event descriptions, and other important attributes. According to these event records, we can segment a whole game into multiple possessions.

4.2 Playing Style Characterization

To incorporate essential game context, we propose a representation learning method based on a Word2Vec model, CBOW [33,34], to characterize players' playing style for describing their decision-making in various contexts. The core idea of Word2Vec is that a word's meaning is determined by its context. It predicts the center word using surrounding words and represents its meaning with an embedding vector. Similarly, we assume a player's style is determined by the decision-making context, since players with similar styles often make the same decisions in similar contexts. Thus, our model uses decision context to predict the corresponding player (Fig. 1). Players making similar decisions in analogous contexts have higher predicted probabilities. Their embedding vectors (Fig. 1(C)) are close in vector space. Despite the existence of advanced pre-trained embedding models (e.g., BERT [16]), considering the current limited dataset, computational limitations, and the requirement for real-time analysis, Word2Vec is an appropriate option. The details are as follows.

Context Vector Formulation. First, we construct a *context vector* to describe the game context in which the player makes a decision (Fig. 1(A)). Following expert advice, we identified 27 crucial features from key moments when the ball handler decides to take a shot, give a pass, or start a dribble. These features are divided into 4 groups: *time* (3), *locations* (22), *score difference* (1), and *decision type* (1).

- *Time* records when the decision is made, including the *period*, *game clock*, and *shot clock*.
- *Locations* contain the X, Y coordinates of the ball handler, the ball, four off-ball offensive players, and five defensive players.
- *Score difference* is calculated by subtracting the defensive team's scores from the offensive team's scores.
- *Decision type* indicates whether the ball handler decides to shoot, pass, or dribble, which is marked by -1, 0, or 1, respectively.

The ball handler's identity serves as the label for the *context vector*. To ensure each player has an adequate number of *context vectors*, we only consider players who have played more than 96 minutes on the court. In our dataset, 336 players meet the requirements.

Model Architecture. Given a *context vector*, the model is designed to predict which ball handler makes the decision in such a game context. In the prediction process, the input *context vector* is first converted into a 300-dimensional hidden vector through normalization and linear transformation. The dot product between this hidden vector and the weight matrix in the output layer is then computed.

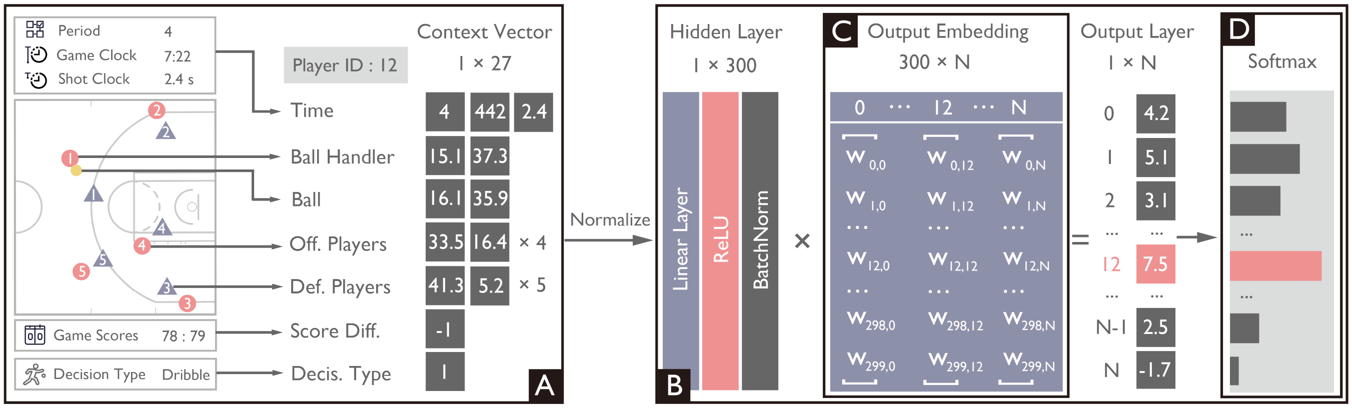

Figure 1: A model for characterizing basketball players' decisions. (A) The construction of a ball handler's context vector, which describes the specific game context when she/he starts a dribble. (B) A predictive model based on the structure of Word2Vec (CBOW) [33]. (C) The latent representation of ball handlers' playing styles in the game context. (D) The prediction of a player's identity given the anonymous context vector.

Finally, the softmax function is applied to generate the probability distribution of all possible ball handlers (Fig. 1(D)).

To reduce the training time, we follow [34] and adopt the Negative Sampling strategy rather than direct optimization methods. For each *context vector*, v_c, we generate negative samples by randomly choosing 5 players except the target ball handler. The sampling process is based on how often the player appears in all *context vectors*. Finally, the objective function is defined as follows:

$$F = \log \sigma(v_h^\top v_{w_O}) + \sum_{w_i \in NEG(w_O)}^{5} \log \sigma(-v_h^\top v_{w_i}) \tag{1}$$

where v_h denotes the transformed hidden vector of v_c, v_{w_O} denotes the weight vector of the actual ball handler, and v_{w_i} denotes the weight vector of the i-th negative player. The training goal is to maximize the value of the above objective function. In other words, the predicted probability of the target ball handler should be as high as possible, while those of the negative players should be as low as possible. Following this pipeline, we train the model using the Adam optimizer with a learning rate of $1e^{-3}$ for 2000 epochs. Finally, each weight vector in the output layer represents the corresponding player's playing style. The similarity between two players' decision-making can be measured by the cosine distance.

4.3 Evaluation

Player Identification. Our model is supposed to identify the ball handler based on the game context in which they make decisions. Inspired by the validation method in Player Vectors [14], which analyzes soccer players' playing styles, we use the accuracy of player identification to verify the model's effectiveness. First, we train the model on half of the games and use the other half for testing. Second, we collect a particular player's *context vectors* from the test data and input them into the model. Third, we add all the output vectors together and obtain the sum of probability values for each player. Finally, we sort these values in descending order and record whether the player appears at the top-1, top-3, top-5, or top-10 results. The above steps are repeated for all 336 players, and the number of players appearing in the top ranks serves as a metric for the identification task. However, there are inherent differences in game contexts between basketball and soccer. In addition, our model incorporates more comprehensive information, such as defenders' positions and game time, instead of relying on match event streams like Player Vectors. A simple comparison experiment might be unfair. Thus, we present only our model's experimental results in Table 1.

Table 1: Experimental Result of the Ball Handler Identification Task

	Top-1	Top-3	Top-5	Top-10
Our Model	47.62%	70.54%	79.76%	85.71%

Decision-Making Similarity. The primary goal of this model is to help find players who make decisions similarly. To validate this supposition, we project players' embedding vectors into a two-dimensional plane by PCA dimensionality reduction (Fig. 2). The color represents the player's playing position, where red indicates guards, yellow indicates centers, and blue indicates forwards. From the figure, we can see that the red and yellow points are clearly clustered, while some of the blue points span between these two categories of points. We speculate that this phenomenon comes from the blurring of the forward's function in modern basketball. Since players in the same position naturally have similar decision-making tendencies, we believe that this two-dimensional projection reflects our model's ability to assess the similarity of players' decisions.

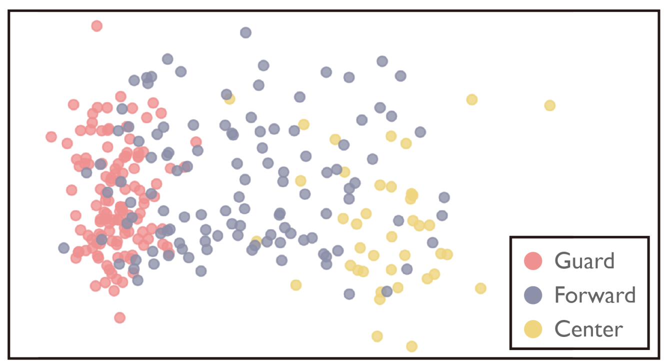

Figure 2: The 2D projection of player embedding vectors.

5 Sketch-based Search

As stated in Section 3, our system should efficiently search for the game situations that match the user sketch. To search through this huge tracking data, we use a method based on spatial indexing, which includes three steps: *simplifying*, *indexing*, and *filtering*.

5.1 Simplifying

We simplify trajectories using partitioning and compression (Fig.3).

Preprocessing. In the user sketch and tracking data, the ball or the player trajectory is represented as a sequence of frames $Tr = \{f_0, f_1, ..., f_n\}$. A frame is denoted by a tuple $f = (t, x, y)$, where t denotes the timestamp and x, y denote the coordinates. To describe the situation more specifically, we add an attribute, ball status, into the frame of the ball trajectory (Fig.3(B)). According to the event description and its timestamp, the ball status is first marked as *shot* or *out-of-control* after a specific event (e.g., *shoot, foul, turnover*). For unmarked frames, we calculate the distance between the ball and the nearest offensive player, based on which each frame is classified as *dribbled* or *passed*. In the end, we get the enhanced frame $f' = (t, x, y, s)$ for the ball trajectory, where s denotes the current ball status.

Spatial Transformation. To simplify spatial information, we refer to [37] and partition the half court into 21 functional divisions

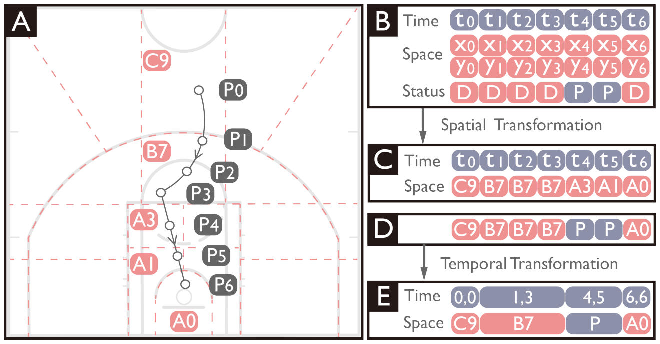

Figure 3: The simplifying process of an example ball trajectory. (A) shows the 21 divisions of a half court and the ball trajectory. The red label is used to identify the division, such as C9 and B7. (B), (C), (D), and (E) show the transformation process of the ball trajectory. D and P in (B) denote *dribbled* and *passed*, respectively.

(Fig.3(A)). The coordinates in each frame are then replaced with the corresponding division label (Fig.3(C)). Considering that the division label may be unstable when the ball or player is at the division boundary, we detect and replace such labels with consistent ones. Moreover, the division label for the ball trajectories is further replaced by P if the current ball status is *passed*, as experts are more concerned with the start and end location rather than the ball's path. Eventually, we get a sequence of division labels, also called the spatial sequence (Fig.3(D)), to represent the spatial information of a trajectory.

Temporal Transformation. To simplify temporal information, we reduce the trajectory size according to its spatial sequence (Fig.3(E)). First, we replace the timestamp of each frame with its index in the raw trajectory. Second, if there are identical division labels consecutively in the spatial sequence, we compress them into a single label. Third, we use an index tuple (i_s, i_e) to record the time for the compressed label, where i_s denotes the start index of the label and i_e denotes the end index. Through the above process, trajectories in the tracking data can be reduced to an average of 10% of their original length.

5.2 Indexing

In the indexing process, we construct a prefix tree based on the simplified sequences of the ball in the tracking data.

Prefix Tree Construction. The prefix tree is a data structure for efficient information retrieval, in which each node represents a character in a word. Following this idea, our tree node represents a division label in a spatial sequence. In addition, the node also contains an index list to find the game situations from the tracking data.

First, for each possession in the database, we obtain the spatial sequence of the ball and generate its subsequences. The length of subsequences ranges from 2 to the maximum length due to sequence variations. For example, given a sequence $S = (C9, B7, P)$, we would get the subsequences $(C9, B7), (B7, P), (C9, B7, P)$. Next, subsequences of different lengths are inserted into the prefix tree one by one. When a subsequence ends at a certain tree node, a pointer to the corresponding situation is added to the node's index list. Finally, we construct an efficient search tree to find the situations that contain similar ball movements. The time complexity of searching a ball sequence in the tree is $O(N)$, where N denotes its length.

Based on the prefix tree, we can extract the ball sequence from the user sketch and take it as the search key, thereby enabling the sketch-based search for game situations.

5.3 Filtering

We filter the search results through a sequence matching strategy.

Filtering by Player Sequences. The user sketch usually involves trajectories of multiple players, making multi-trajectory search necessary. Therefore, we perform an additional filtering step after obtaining the situations from the prefix tree. First, we extract the player sequences from the sketch and use them as filters. Next, for a retrieved situation, we find all player sequences that contain the same spatial sequence as the filters. The temporal sequences are then used to locate and compare the start and end times of these spatial sequences. If each

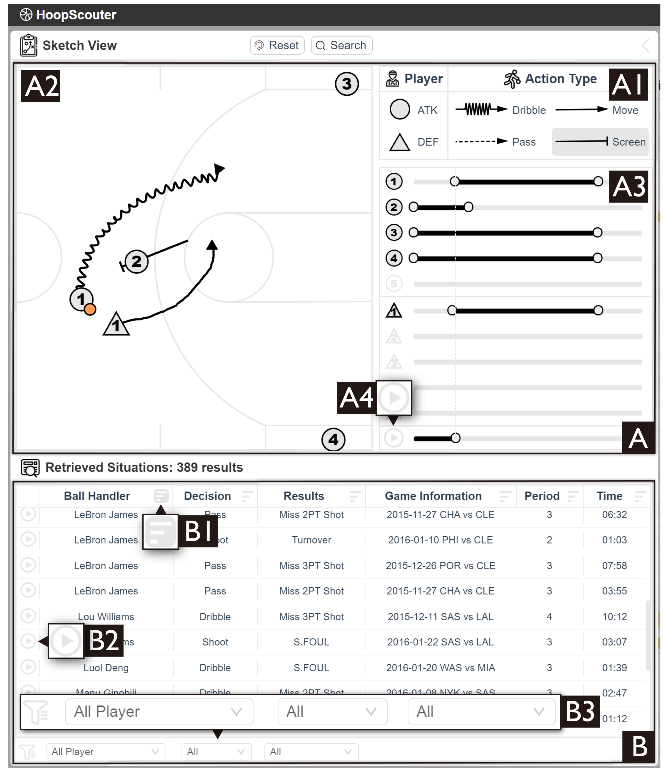

Figure 4: The sketch view of HoopScouter. (A) The Sketch Panel enables users to create a query for a specific game situation. (B) The Situation Table lists the situations retrieved from the tracking data.

filter matches at least one player sequence within the corresponding time span, the situation would be regarded as a valid search result. In the end, we can filter to get the game situations that also contain similar player movements.

Overall, HoopScouter integrates this sketch-based method to search for a specific situation. In our case, users can search through more than 1.2 million trajectories with a response time of about 0.1 seconds. Besides, the ball handlers' decisions made at the end of the situation can be inferred from the complete trajectories in the search results.

6 Visual Design

We develop a visual analytics system, HoopScouter, to meet the requirements in Section 3.2. The system comprises four views: the sketch view (**L1, L2**), the decision view (**S1, S2, and C1, C2**), the player view (**C1**), and the factor view (**C2**). Starting from the sketch view (Fig. 4), users can create a query by depicting a game situation of interest. Then in the decision view (Fig. 5(B)), they can observe the ball handlers' decisions obtained from the search results. While analyzing the results, users may modify the player groups through the player view (Fig. 5(C)) and compare different players' decisions. Finally, in the factor view (Fig. 5(D)), they can inspect how a contextual factor influences the player group's decision-making.

6.1 Sketch View

The sketch view (Fig. 4) displays both the input query and the search results. It consists of two parts: (1) the **Sketch Panel** supports the visual depiction of a specific situation and generates a query for the situation (**L1, L2**); (2) the **Situation Table** provides basic information about the situations retrieved from the tracking data (**L2**).

Sketch Panel (Fig. 4(A))**.** The panel is composed of three components: a tool palette, a whiteboard, and a timeline chart.

- *Tool palette* (Fig. 4(A1)). Coaches typically depict game situations via the play diagram [58]. Thus, we follow this approach to design the tool palette, where icon shapes (circle and triangle) indicate player types (attacker and defender) and line types denote action types (e.g., dribble and screen). Users can select buttons in the tool

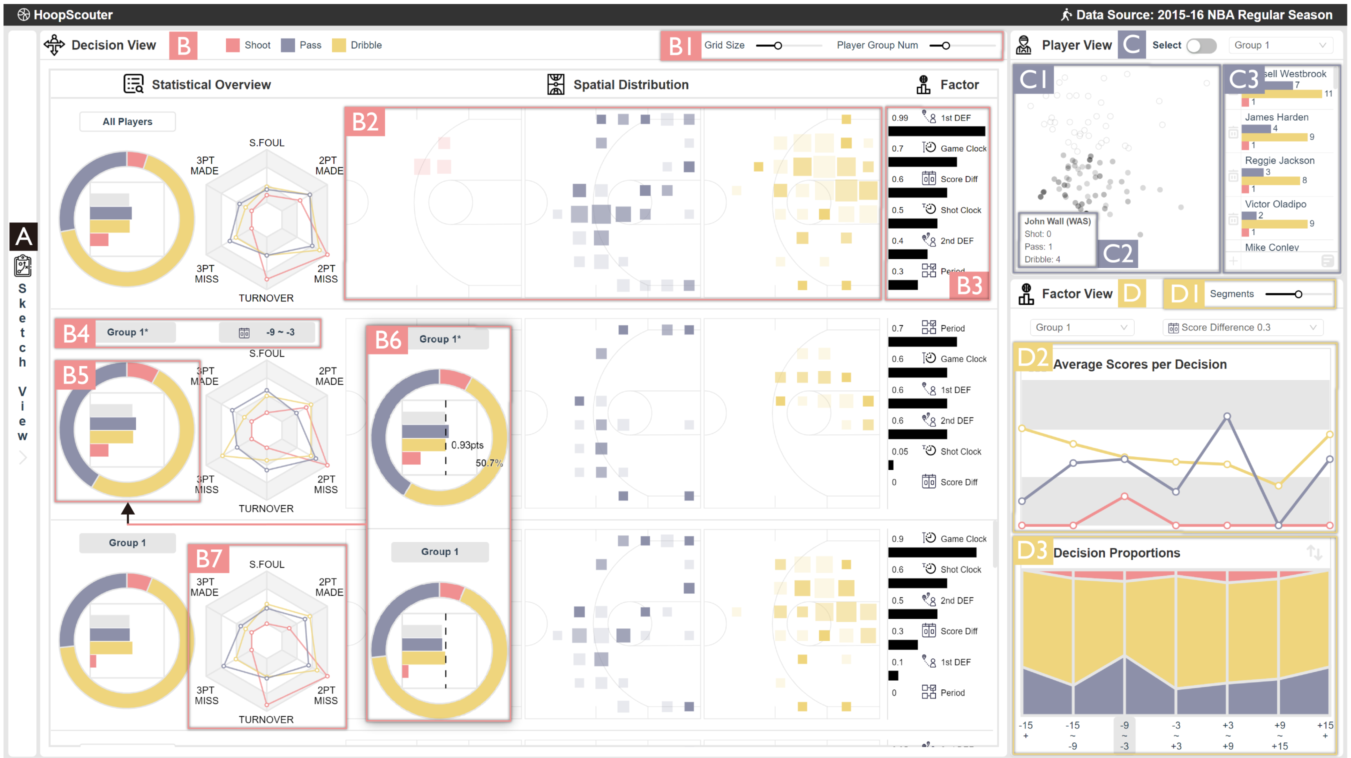

Figure 5: The interface of HoopScouter. (A) The sketch view appears in a collapsed state and expands when the title is clicked. (B) The decision view compares decisions across player groups. (C) The player view shows a group's playing styles and decisions. (D) The factor view highlights how contextual factors, like score difference, influence decisions.

palette to add elements to the whiteboard, such as a trajectory of player screening (Fig. 4(A1)).

- *Whiteboard* (Fig. 4(A2)). On the whiteboard, the user-drawn trajectory shows the path of a player on the basketball court. The player icon reflects her/his location at a specific time point, while the orange circle indicates the ball's location. The entire ball trajectory can be inferred from the trajectories of successive dribbles or passes. For instance, the ball trajectory in Fig. 4(A2) is the same as the dribbling path of attacker 1.
- *Timeline* (Fig. 4(A3)). Users need temporal information to describe action sequences or movement speed. For example, in Fig. 4(A2), attacker 2 should set the screen before attacker 1 starts to dribble. Hence we apply a timeline, where each black bar corresponds to one player trajectory on the whiteboard. The horizontal axis ranges from 0 to 10 seconds, and the endpoints of the bar indicate when the player starts or finishes moving.

Situation Table (Fig. 4(B))**.** The table lists all the search results after executing a query via the method in Section 5. Each table row represents a retrieved game situation, and each column offers essential information about the situation, such as the ball handler. Besides, sorting buttons (Fig. 4(B1)) and filtering methods (Fig. 4(B3)) are provided to facilitate the exploration of the search results.

Interactions. The main interactions are presented below.

- *Draw player trajectories.* At first, users should select an element in the tool palette. Then they can click on the whiteboard to set a new player icon, or drag and drop an existing icon to draw a trajectory. During dragging, the panel captures the mouse location in real time and generates a sequence of points to record the trajectory. The time duration of the trajectory would initially be inferred from its length and plotted in the timeline.
- *Adjust time information.* The timeline allows users to explicitly determine when and how fast the player moves in the sketched situation. Users can drag the endpoint of the bar to set when the player starts or stops moving. As it changes the bar length, the point sequence of the trajectory would be upsampled or downsampled accordingly, thus changing the player speed.
- *Display player movements.* We add play buttons in both the timeline (Fig. 4(A4)) and the Situation Table (Fig. 4(B2)). Users can click it to see an animation of the players moving on the whiteboard, confirming whether the game situation satisfies their expectations.

6.2 Decision View

The decision view (Fig. 5(B)) provides multiple perspectives to observe the ball handlers' decisions obtained from the search results: (1) the **Statistical Overview** presents key indicators of the decisions (**S1**); (2) the **Spatial Distribution** reveals the target locations and corresponding effects of the decisions (**S2**); (3) the **Factor Bars** reflect the influences of contextual factors on players' decision-making (**C2**). Each row shows the decisions made by a group of players, which supports comparative analysis between players (**C1**).

Statistical Overview. First, we use labels to categorize the decisions of different groups. For instance, the labels in Fig. 5(B4) indicate that this row presents the decisions of player group 1 when the score difference is between -9 and -3 points. Second, we combine two charts to visualize statistical indicators (Fig. 5(B5)). The donut chart shows the proportion of different decisions, where the color hue indicates the decision type (red for shooting, blue for passing, and yellow for dribbling). Meanwhile, the bar chart shows the average points scored by a shot, pass, and dribble. The gray bar indicates the average points per decision and serves as the baseline. Third, we add a radar chart to display the concrete outcomes of the decisions (Fig. 5(B7)). Each axis represents a type of event that may occur at the end of the possession, such as a shooting foul. The distance from the origin represents how often the event occurs.

Spatial Distribution. Inspired by the design of CourtVision [21], we employ a grid-based visualization to present spatial information. As shown in Fig. 5(B2), the three basketball courts sequentially show where the ball handlers shoot at, pass or dribble the ball to. The court is divided into an $N \times N$ grid, where each grid cell contains a square to represent the decisions. The size of the square indicates the number of decisions targeting this location, and the opacity encodes the average scoring rate of these decisions.

Factor Bars. We consulted experts and picked six contextual factors that may influence the ball handlers' decision-making, including

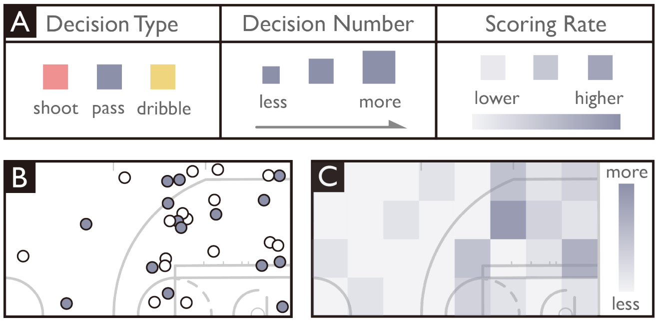

Figure 6: The visual design of the Spatial Distribution and two alternatives. (A) shows the visual encoding of a grid cell.

period, game clock, shot clock, score difference, and distances to the nearest and next-nearest defender. To assess how decision are associated with these factors, we conduct chi-square tests and calculate the p-value based on the retrieved decisions. The "1 minus p-value" is then presented in the bar chart (Fig. 5(B3)). The longer the bar, the more likely the factor is to affect the players' decision-making.

Interactions. This view supports interactions as follows.

- *Show decision statistics.* In Statistical Overview, users can hover over a bar to view specific scores and percentages (Fig. 5(B6)), with a dashed line aiding comparisons across player groups.
- *Adjust grid size.* The left slider in Fig. 5(B1) enables users to set the grid size of the court. Users can thus view the distribution of decisions at different levels of detail.
- *Adjust group number.* K-Means clustering is applied to adjust the number of player groups based on the player embeddings, which is described in Section 4. The right slider in Fig. 5(B1) allows users to set the number of player groups in the decision view. New player groups are initialized using K-Means clustering (initial $k = 3$) based on decision categories.

Justification. We justify two key design choices in this view.

- *Statistical Overview.* We explored visualizations that show proportions, such as pie charts, Sankey diagrams, proportional areas, and stacked bar graphs. Among these alternatives, only the stacked bar graph could be compactly combined with a bar chart showing the scores. Nevertheless, such a combination sometimes confused our experts, as the bar length represented both the proportion and score. Therefore, we opted for the donut chart, as its inner radius would not seriously impair the perception of proportions [8].
- *Spatial Distribution.* We also offered two design alternatives (scatter plot and heatmap) to present spatial information. For instance, in Fig. 6(B), each dot marks the target location of a pass. Hollow and solid dots denote unsuccessful and successful outcomes, respectively. In Fig. 6(C), the opacity of each cell encodes the density of the dots. However, our experts found it difficult to discern scoring rates, especially in dense areas. We modified a grid-based design [21] to support the analysis. We reserved color hue for decision types to maintain consistency in visual encoding, and used opacity to represent scoring rates considering its strong interaction effects with color hue [36].

6.3 Player View

After determining the number of player groups displayed in the decision view, the player view (Fig. 5(C)) helps users explore and modify the players in each group (**C1**). Users can select a group to view their playing styles and decisions.

Through the method in Section 4, we can use an embedding vector to represent the ball handler's playing style. Thus, the scatter plot (Fig. 5(C1)) presents a 2D PCA projection of the player embeddings. In other words, players who tend to make similar decisions are clustered together. In this chart, each dot represents a ball handler that appears in the search result. Hollow dots indicate players who do not belong to the selected group, whereas solid dots do the opposite. The opacity of solid dots encodes the number of decisions made by the player. The bar chart (Fig. 5(C3)) displays the list of players in the selected group. Each row displays the number of decisions per type for each player. To facilitate exploration, users can sort these rows by player name or by decision count.

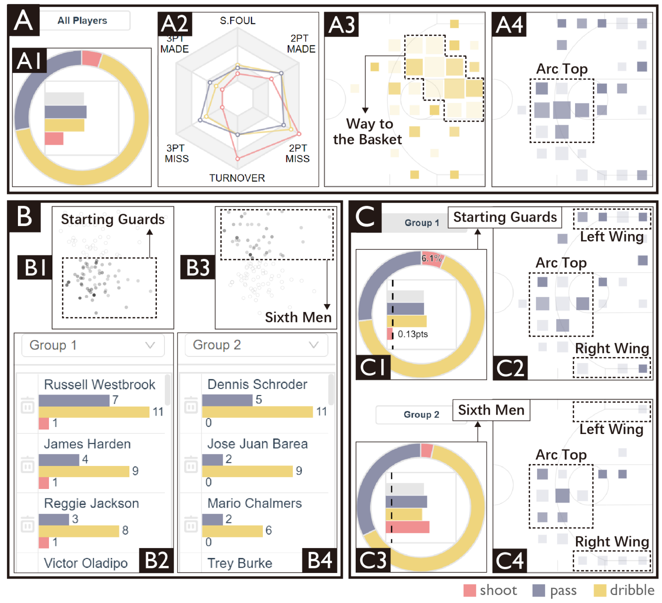

Figure 7: The analysis process of the first case. (A) presents the statistical indicators and spatial information of all player decisions. (B) presents two player groups. (C) presents the comparison between the starting guards and sixth men.

Interactions. This view includes interactions as follows.

- *Show player information.* Users can hover over a dot to see a tooltip containing the player information (Fig. 5(C2)), such as her/his name, team, and decisions made.
- *Select players.* By clicking the button, users can switch to the select mode and draw a rectangular selection around dots. The players in the current group are replaced with the selected ones.
- *Modify players.* The bar chart allows users to modify the player list in the current group. Users can remove players by clicking the delete button on the left of each row or add players by entering their names at the bottom. When a group is modified, the decision view and factor view are updated accordingly.

6.4 Factor View

The factor view (Fig. 5(D)) illustrates how a group of players is affected by a contextual factor (**C2**). It contains two parts: (1) the line chart (Fig. 5(D2)) shows how the average score for each decision type varies with a particular factor. Users can hover to see specific scores. (2) the percentage stacked area chart (Fig. 5(D3)) reflects changes in decision proportions. Users can click the button in the top right corner to reorder the stacks. Both charts share the same factor axis, uniformly divided into multiple segments.

Interactions. The primary interactions are shown below.

- *Select a group/factor.* By selecting from the drop-down lists above the line chart, users can determine the group or factor to be displayed in the factor view.
- *Adjust segment number.* The slider (Fig. 5(D1)) enables users to change the number of segments for the axis.
- *Show decision details.* To view the decision details within a specific segment, users can click the segment label to add a row in the decision view. For example, the row (Fig. 5(B4)) is added after clicking the label of "-9 $\sim$ -3".

7 System Evaluation

We conducted case studies with three domain experts, E1, E2 and E3. After being introduced to the system's visualizations and interactions, they freely sketched situations and conducted decision-making analyses for 30 minutes, followed by a feedback interview.

7.1 Decision-making when running pick-and-rolls.

The first case focuses on the ball handlers' decision-making when executing an important play, the pick-and-roll [64]. The play usually involves three players, with one screener blocking the path of the defender and creating space for the ball handler. At this moment, the ball handler's decision essentially determines the success or failure of the attack. E1 began with the sketch view to depict a typical situation of pick-and-rolls (Fig.4(A2)). After sketching the paths on the whiteboard, E1 adjusted the Gantt chart (Fig.4(A3)) and watched an animation to ensure that attacker 2 set a screen before attacker 1 started dribbling. Then E1 created a query and obtained a total of 389 similar situations. He randomly played the animations of several situations and confirmed that the search results satisfied his expectations. All experts thought highly of the Sketch Panel, with one commenting that "*it provides an intuitive way to quickly find the game situations we want to analyze.*"

Insight 1: Jump shots were inefficient decisions. The experts turned to the decision view for valuable insights. When examining the Statistical Overview, they found that a jump shot made much fewer points than a pass or dribble (Fig.7(A1)), as most shots were missed or led to a turnover (Fig.7(A2)). E1 approved that the overall offensive efficiency of such shooting was fairly low. He then went back to the Situation Table and browsed the ball handlers who chose to shoot when running pick-and-rolls. Among these players, Trey Burke was the only one who chose to shoot more than once. E1 thus suggested that "*he may need to change such decision-making habits to improve his offensive performance.*"

Insight 2: The ball handlers were prone to dribble to the basket or pass to the screener. The experts continued to investigate the Spatial Distribution. At first, they found many yellow squares on the way to the basket (Fig.7(A3)). As dribbling was the primary decision in the donut chart (Fig.7(A1)), they deduced that most ball handlers would continue dribbling to the basket. Meanwhile, a few large blue squares appeared at the arc top (Fig.7(A4)). This meant that the ball handlers' main passing target was the screener (i.e., the player who set the screen). They explained, "*the defensive center usually filled in to defend the ball handler, thus the screener gaining the open shot.*"

Insight 3: The starting guards and sixth men differed considerably in their decision-making. The two groups of players shown in the decision view were initialized from the K-Means clustering by default (Fig.7(B1, B3)). HoopScouter allows the experts to manually adjust the player groups and further explored the groups in the player view. To verify that players within the same group had similar playing styles, E1 browsed the player list for each group (Fig.7(B2, B4)). Then he found the first group was mainly starting guards (i.e., players who control the ball in the starting lineup), and the second group was mainly sixth men (i.e., players who control the ball in the substitute lineup). After removing several outliers for each group, the experts returned to the decision view and compared the decision-making of the two player groups. By hovering on the bars, they realized that the starting guards were more likely to shoot the ball, but scored far fewer points for it than the sixth men (Fig.7(C1, C3)). From the Spatial Distribution, E2 noticed that the sixth men barely passed the ball to the left wing (Fig.7(C4)), while the starting guards nearly achieved a balance in both wings (Fig.7(C2)). Meanwhile, E1 speculated that cooperating with the screener was one of the most effective ways for the sixth men, as the largest square at the arc top showed a high opacity (Fig.7(C4)). After a brief discussion, the experts provided a targeted defense strategy for the sixth men. "*The defenders in the left wing should participate in the defense of the sixth men, thereby reducing the burden on the defensive center.*"

7.2 Decision-making when running high post offense.

The second case involves another common play, the high post offense [45]. This play usually starts with a player holding the ball in the high post (i.e., the area near the free throw line). The ball handler should be flexible in their decision-making to lead the attack.

We invited E3 to conduct the case study. With the Sketch Panel, E3 depicted a usual situation of the high post offense (Fig.8(A1, A2)), namely, attacker 2 received the ball at the high post with another off-ball attacker cutting to the basket. After watching the animation for verification, E3 clicked the search button. Then, a total of 413 similar situations were retrieved and listed in the Situation Table.

Insight 1: Giving a pass was the most efficient decision for the team offense. E3 investigated the decision view to gain insights from the decision-making of the ball handlers. As shown in Fig.8(B1), even though dribbling was the primary decision of the ball handler, passing was the decision with the highest offensive efficiency. E3 learned the reason from the radar chart (Fig.8(B2)): passing was more likely to lead a made two-point shot compared to shooting and dribbling. Besides, similar to the **Insight 1** of the first case, shooting was still the most inefficient decision.

Insight 2: Most playmakers chose to dribble back to the three-point line. E3 then explored the player view to see how the ball handlers were grouped. The scatter plot was divided into two parts by default, with players in group 1 appearing in the left half (Fig.8(C1)). From the player list of group 1 (Fig.8(C3)), he realized that this group contained a lot of playmakers (i.e., players who lead the team offense). To examine how the playmakers handled the ball in the high post offense, E3 went back to the decision view. The donut chart (Fig.8(D1)) showed that the playmakers would choose to dribble almost three out of four times. Additionally, the biggest yellow square emerged outside the three-point line (Fig.8(D2)), which means that playmakers mostly dribbled back when getting the ball in the high post. E3 thus inferred that "*these playmakers dribbled back to the three-point line to call for a screen or reorganize the attack, as the high post might be an uncomfortable position for them to create scoring opportunities.*"

Insight 3: The post players played a significant role in the ball movement. The players in group 2 were distributed in the right half of the scatter plot (Fig.8(C2)). By browsing the player list (Fig.8(C4)), E3 found that most of them were post players (i.e., players who operate near the baseline). To compare with the playmakers, E3 hovered the bar in the Statistical Overview. With the assistance of the dashed line, he deduced that passing was not only the primary decision for the post players, but outperformed all types of decisions of the playmakers (Fig.8(D1, E1)). In terms of the Spatial Distribution, nearly no yellow squares emerged outside the three-point line (Fig.8(E2)). It implied that the basic target of a post player dribbling was to move towards the basket. In addition, the passing targets were distributed over the court, with the biggest blue square showing up at the arc top (Fig.8(E3)). E3 suggested, "*The post players often served as the pivot for the ball movement, so at this moment, the defenders should be alert to the off-ball attackers in good scoring positions.*"

Insight 4: The decision-making of post players was most affected by the score difference. According to the bar chart (Fig.8(F1)), the score difference was the factor that had the greatest impact on the decisions of post players. E3 thus turned to the factor view and found that the possibility of a post player passing the ball was gradually decreasing as the score went from behind to ahead (Fig.8(F2)). Meanwhile, the possibility of a jump shot was increasing. As the jump shot was not a good offensive decision overall (**Insight 1**), such a tendency might give the opponent a chance to tie the scores.

7.3 Expert Feedback

After the case studies, we conducted interviews with the experts and summarized their feedback on HoopScouter.

Usability. In general, the experts were satisfied with the usability of HoopScouter for three main reasons. First, the sketch view effectively helps the experts to find the game situations of interest. E1 stated that "*the search results generally match my sketch.*" E2 also commented that "*the novel panel is helpful to describe a specific situation,*" and "*the animation can assist teenage players in understanding complex tactics.*" Second, the default grouping of ball handlers facilitates to uncover the typical decision-making patterns of different players. E3 mentioned that "*the auto-generated player groupings agree with my empirical knowledge,*" and "*the decision view shows the difference in decision-making between players with specific roles.*" Third, the factor view supports the discovery of latent associations between game contexts and ball handlers' decision-making. E1 and E2 commented that "*the influence of contextual factors is intuitively explained by the change of decision-making tendency and effects.*"

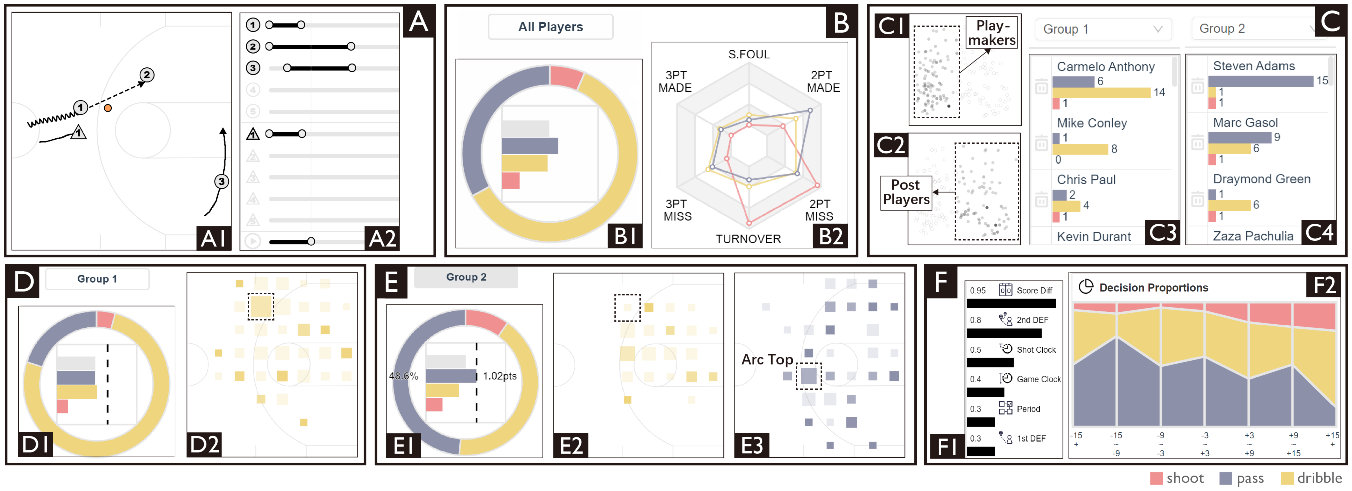

Figure 8: The analysis process of the second case. (A) presents the user sketch of the high post offense, which served as the search input. (B) presents the statistical indicators of all player decisions. (C) presents two player groups. (D) and (E) present the comparison between the playmakers and post players. (F) presents how the post players were affected by the score difference.

Suggestions. The experts also offered some suggestions for future improvement. First, standard templates can be added to simplify the depiction of the game situations. E1 suggested, "*Some common situations in games can be used as templates, then I can quickly depict a situation with just a few modifications based on the template.*" Second, filtering based on target locations can be enabled for the search results. E3 told us, "*We can conduct a more detailed analysis if the system supports location-based filtering, such as the situations that the ball handler dribbled back to the three-point line.*"

8 Discussion

Lessons Learned. We have learned two valuable lessons from the design study [32, 48]. First, time information can be necessary when presenting the trajectory data. In addition to the players' movement paths, their action sequence and movement speed are essential for describing a specific situation in basketball games. Therefore, in the Sketch Panel, we integrate the time information of player trajectories into the visual depiction. Additional time information provided by the user can also improve the accuracy of the search results. Second, representation learning methods can accelerate the data exploration and pattern recognition. In the player view, we provide a projection of the player embeddings, which are generated by characterizing the ball handlers' playing styles. Utilizing the projection, users can quickly obtain a group of players who tend to make similar decisions at the end of a specific game situation. In the future, incorporating methods (eg., elbow method [47]) for the k value selection could reduce the burden on users from directly examining the player list.

Significance. Decision-making is an key skill for players in competitive sports like basketball [10] and soccer [44]. Thus, we introduce a visual analytics system, HoopScouter, to facilitate the investigation and understanding of the ball handlers' decision-making in basketball games. With HoopScouter, coaches can retrieve historical situations that match their sketches, identify decision-making patterns for a group of players, and compare decisions between players or under different game contexts. Besides, the experts believe our player embeddings would be valuable for scouting or drafting a player. It may assist them in discovering players with similar playing styles, thus avoiding the burden of long-term observations for these players.

Scalability. In this study, we use a single-season tracking dataset to evaluate the effectiveness of HoopScouter. Players' decisions are often shaped by their team's tactical strategies (e.g., ball handlers in teams that emphasize ball-sharing tend to pass more frequently). Since it's common for players to transfer across seasons, focusing on a single season with a stable lineup ensures more accurate and consistent analysis. When it comes to data scalability, HoopScouter's data retrieval and visualization are not limited by dataset size. Interactions such as filtering and sorting help users quickly find specific plays of interest. However, these goal-oriented interactions are less effective for general browsing. For larger datasets, features like pagination and icon-based visualizations can further enhance usability.

Generalizability. Strategically planned movements are frequently executed in many other competitive sports, such as soccer, rugby, and ice hockey. After a reasonable modification, our Sketch Panel and search engine can be applied in these sports to search for similar game situations. Furthermore, players' decision-making also plays a significant role in these sports. It is often associated with the game contexts, such as the locations of players, the score difference, and the defense. Given the similarities between these sports, our representation learning method is applicable to characterize the decision-making of players in these sports as well.

Limitations. Our work has two limitations. The first one lies in the insufficient consideration of playing skills when characterizing the ball handlers' decision-making, such as shot fakes, footwork, and dribbling skills. For instance, the ball handler may make a fake shot to get past the defender. However, it is difficult to capture playing skills from tracking data due to the lack of information on the players' body movements. As body movement data might be available in the future, we have plans to further improve our model by incorporating the playing skills. The second limitation lies in the interactions of drawing the defenders' paths. In most cases, the defenders would follow a defensive strategy, such as zone defense and man-to-man defense, and move to catch up with the offensive players. Due to the regularity of the defenders' movements, it is not very meaningful to draw the paths of the defenders one by one. In the future, we plan to replace such tedious interactions by using a filter based on the defensive strategy, thus simplifying the steps for users to create a sketch-based query.

9 Conclusion

We present HoopScouter, an interactive visual analytics system for exploring and understanding basketball ball handlers' decision-making. In collaboration with experts, we define the analysis workflow and key domain problems. A sketch-based search retrieves player decisions for specific situations, visualized with tailored designs. To compare decisions across players, we use representation learning to embed playing styles. Usability and effectiveness are validated through two case studies and expert interviews. In the future, we plan to extend HoopScouter to examine the decision-making of off-ball players. It can help experts study the collaborative activities in basketball games and thus develop more effective offensive tactics.

Acknowledgments

The work was supported by NSFC (U22A2032) and the Collaborative Innovation Center of Artificial Intelligence by MOE and Zhejiang Provincial Government (ZJU).

REFERENCES

[1] SportVU. https://www.stats.com/sportvu-basketball/, 2010.

[2] S. Afzal, S. Ghani, H. C. Jenkins-Smith, D. S. Ebert, M. Hadwiger, and I. Hoteit. A visual analytics based decision making environment for COVID-19 modeling and visualization. In *2020 IEEE Visualization Conference (VIS)*, pp. 86–90, 2020.

[3] G. Andrienko, N. Andrienko, G. Anzer, P. Bauer, G. Budziak, G. Fuchs, D. Hecker, H. Weber, and S. Wrobel. Constructing spaces and times for tactical analysis in football. *IEEE Transactions on Visualization and Computer Graphics*, 27(4):2280–2297, 2019.

[4] G. Andrienko, N. Andrienko, and P. Jankowski. Building spatial decision support tools for individuals and groups. *Journal of Decision Systems*, 12(2):193–208, 2003.

[5] J. Bautista and G. Carenini. An integrated task-based framework for the design and evaluation of visualizations to support preferential choice. In *Proceedings of the working conference on Advanced visual interfaces*, pp. 217–224, 2006.

[6] P. Beshai. Buckets: Visualizing NBA shot data. *University of British Columbia, published Dec*, 2014.

[7] A. Bialkowski, P. Lucey, P. Carr, Y. Yue, S. Sridharan, and I. Matthews. Identifying team style in soccer using formations learned from spatiotemporal tracking data. In *2014 IEEE International Conference on Data Mining Workshop*, pp. 9–14, 2014.

[8] X. Cai, K. Efstathiou, X. Xie, Y. Wu, Y. Shi, and L. Yu. A study of the effect of doughnut chart parameters on proportion estimation accuracy. In *Computer Graphics Forum*, vol. 37, pp. 300–312, 2018.

[9] G. Carenini and J. Loyd. Valuecharts: analyzing linear models expressing preferences and evaluations. In *Proceedings of the working conference on Advanced visual interfaces*, pp. 150–157, 2004.

[10] D. Cervone, A. D'Amour, L. Bornn, and K. Goldsberry. Pointwise: Predicting points and valuing decisions in real time with nba optical tracking data. In *Proceedings of the 8th MIT Sloan Sports Analytics Conference, Boston, MA, USA*, vol. 28, p. 3, 2014.

[11] C.-H. Chen, T.-L. Liu, Y.-S. Wang, H.-K. Chu, N. C. Tang, and H.-Y. M. Liao. Spatio-temporal learning of basketball offensive strategies. In *Proceedings of the 23rd ACM international conference on Multimedia*, pp. 1123–1126, 2015.

[12] C.-Y. Chen, W. Lai, H.-Y. Hsieh, W.-H. Zheng, Y.-S. Wang, and J.-H. Chuang. Generating defensive plays in basketball games. In *Proceedings of the 26th ACM international conference on Multimedia*, pp. 1580–1588, 2018.

[13] W. Chen, T. Lao, J. Xia, X. Huang, B. Zhu, W. Hu, and H. Guan. GameFlow: Narrative visualization of NBA basketball games. *IEEE Transactions on Multimedia*, 18(11):2247–2256, 2016.

[14] T. Decroos and J. Davis. Player vectors: Characterizing soccer players' playing style from match event streams. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 569–584, 2019.

[15] D. Demaj. Geovisualizing spatio-temporal patterns in tennis: An alternative approach to post-match analysis. In *Proceedings of the 26th International Cartographic Conference*, vol. 9, p. 16, 2013.

[16] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *North American Chapter of the Association for Computational Linguistics*, 2019.

[17] M. Du and X. Yuan. A survey of competitive sports data visualization and visual analysis. *Journal of Visualization*, (1):47–67, 2021.

[18] P. Felsen, P. Agrawal, and J. Malik. What will happen next? forecasting player moves in sports videos. In *Proceedings of the IEEE international conference on computer vision*, pp. 3342–3351, 2017.

[19] A. Franks, A. Miller, L. Bornn, and K. Goldsberry. Characterizing the spatial structure of defensive skill in professional basketball. *Annals of Applied Statistics*, 9(1):94–121, 2015.

[20] A. Franks, A. Miller, L. Bornn, and K. Goldsberry. Counterpoints: Advanced defensive metrics for NBA basketball. In *9th Annual MIT Sloan Sports Analytics Conference, Boston, MA*, 2015.

[21] K. Goldsberry. CourtVision: New visual and spatial analytics for the nba. In *Proceedings of MIT Sloan sports analytics conference*, vol. 9, pp. 12–15, 2012.

[22] S. Gratzl, A. Lex, N. Gehlenborg, H. Pfister, and M. Streit. LineUp: Visual analysis of multi-attribute rankings. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2277–2286, 2013.

[23] L. Gyarmati and M. Hefeeda. Analyzing in-game movements of soccer players at scale. *arXiv preprint arXiv:1603.05583*, 2016.

[24] M. Harmon, A. Ebrahimi, P. Lucey, and D. Klabjan. Predicting shot making in basketball learnt from adversarial multiagent trajectories. *arXiv preprint arXiv:1609.04849*, 2016.

[25] H.-Y. Hsieh, C.-Y. Chen, Y.-S. Wang, and J.-H. Chuang. BasketballGAN: Generating basketball play simulation through sketching. In *Proceedings of the 27th ACM International Conference on Multimedia*, pp. 720–728, 2019.

[26] P. A. Legg, D. H. Chung, M. L. Parry, R. Bown, M. W. Jones, I. W. Griffiths, and M. Chen. Transformation of an uncertain video search pipeline to a sketch-based visual analytics loop. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2109–2118, 2013.

[27] T. Lin, Z. Chen, Y. Yang, D. Chiappalupi, J. Beyer, and H. Pfister. The quest for Omnioculars: Embedded visualization for augmenting basketball game viewing experiences. *IEEE Transactions on Visualization and Computer Graphics*, 2022.

[28] D. Liu, D. Weng, Y. Li, J. Bao, Y. Zheng, H. Qu, and Y. Wu. SmartAdp: Visual analytics of large-scale taxi trajectories for selecting billboard locations. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):1–10, 2017.

[29] A. G. Losada, R. Therón, and A. Benito. BKViz: A basketball visual analysis tool. *IEEE Computer Graphics and Applications*, 36(6):58–68, 2016.

[30] R. Metoyer, Q. Zhi, B. Janczuk, and W. Scheirer. Coupling story to visualization: Using textual analysis as a bridge between data and interpretation. In *Proceedings of the 23rd International Conference on Intelligent User Interfaces*, pp. 503–507, 2018.

[31] R. Metulini, M. Manisera, and P. Zuccolotto. Modelling the dynamic pattern of surface area in basketball and its effects on team performance. *Journal of Quantitative Analysis in Sports*, 14(3):117–130, 2018.

[32] M. Meyer and J. Dykes. Criteria for Rigor in Visualization Design Study . *IEEE Transactions on Visualization & Computer Graphics*, 26(01):87–97, 2020.

[33] T. Mikolov, K. Chen, G. Corrado, and J. Dean. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781*, 2013.

[34] T. Mikolov, I. Sutskever, K. Chen, G. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. *arXiv preprint arXiv:1310.4546*, 2013.

[35] T. Munzner. A nested model for visualization design and validation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):921–928, 2009.

[36] T. Munzner. *Visualization Analysis and Design*. CRC Press, 2014.

[37] H. Nunes, X. Iglesias, G. Daza, A. Irurtia, T. Caparrós, and M. T. Anguera. The influence of pick and roll in attacking play in top-level basketball. *Cuadernos de Psicologia Del Deporte*, 16(1):129–142, 2016.

[38] S. Pajer, M. Streit, T. Torsney-Weir, F. Spechtenhauser, T. Möller, and H. Piringer. Weightlifter: Visual weight space exploration for multi-criteria decision making. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):611–620, 2016.

[39] C. Perin, R. Vuillemot, and J.-D. Fekete. SoccerStories: A kick-off for visual soccer analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2506–2515, 2013.

[40] C. Perin, R. Vuillemot, C. D. Stolper, J. T. Stasko, J. Wood, and S. Carpendale. State of the art of sports data visualization. In *Computer Graphics Forum*, vol. 37, pp. 663–686, 2018.

[41] T. Polk, D. Jäckle, J. Häußler, and J. Yang. CourtTime: Generating actionable insights into tennis matches using visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):397–406, 2020.

[42] T. Polk, J. Yang, Y. Hu, and Y. Zhao. TenniVis: Visualization for tennis match analysis. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2339–2348, 2014.

[43] L. Probst, F. Rauschenbach, H. Schuldt, P. Seidenschwarz, and M. Rumo. Integrated real-time data stream analysis and sketch-based video retrieval in team sports. In *Proceedings of International Conference on Big Data*,

pp. 548–555, 2018.

[44] A. Roca, P. R. Ford, and D. Memmert. Creative decision making and visual search behavior in skilled soccer players. *PloS one*, 13(7):e0199381, 2018.

[45] Rookieroad. Basketball high and low post. `https://www.rookieroad.com/basketball/strategy/high-and-low-post-offense/`.

[46] D. Sacha, F. Al-Masoudi, M. Stein, T. Schreck, D. A. Keim, G. Andrienko, and H. Janetzko. Dynamic visual abstraction of soccer movement. In *Computer Graphics Forum*, vol. 36, pp. 305–315, 2017.

[47] T. V. Sai Krishna, A. Yesu Babu, and R. Kiran Kumar. Determination of optimal clusters for a non-hierarchical clustering paradigm k-means algorithm. In *Proceedings of International Conference on Computational Intelligence and Data Engineering*, pp. 301–316, 2018.

[48] M. Sedlmair, M. Meyer, and T. Munzner. Design study methodology: Reflections from the trenches and the stacks. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2431–2440, 2012.

[49] D. Seebacher, T. Polk, H. Janetzko, D. Keim, T. Schreck, and M. Stein. Investigating the Sketchplan: A novel way of identifying tactical behavior in massive soccer datasets. *IEEE Transactions on Visualization and Computer Graphics*, 2021.

[50] T. Seidl, A. Cherukumudi, A. Hartnett, P. Carr, and P. Lucey. Bhostgusters: Realtime interactive play sketching with synthesized nba defenses. In *12th Annual MIT Sloan Sports Analytics Conference*, 2018.

[51] L. Sha, P. Lucey, Y. Yue, X. Wei, J. Hobbs, C. Rohlf, and S. Sridharan. Interactive sports analytics: An intelligent interface for utilizing trajectories for interactive sports play retrieval and analytics. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 25(2):1–32, 2018.

[52] L. Sha, P. Lucey, S. Zheng, T. Kim, Y. Yue, and S. Sridharan. Fine-grained retrieval of sports plays using tree-based alignment of trajectories. *arXiv preprint arXiv:1710.02255*, 2017.

[53] L. Shao, D. Sacha, B. Neldner, M. Stein, and T. Schreck. Visual-interactive search for soccer trajectories to identify interesting game situations. *Electronic Imaging*, 2016(1):1–10, 2016.

[54] A. Sicilia, K. Pelechrinis, and K. Goldsberry. DeepHoops: Evaluating micro-actions in basketball using deep feature representations of spatio-temporal data. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2096–2104, 2019.

[55] R. Sisneros and M. Van Moer. Expanding plus-minus for visual and statistical analysis of NBA box-score data. In *The 1st workshop on sports data visualization*, 2013.

[56] M. Stein, H. Janetzko, A. Lamprecht, T. Breitkreutz, P. Zimmermann, B. Goldlücke, T. Schreck, G. Andrienko, M. Grossniklaus, and D. A. Keim. Bring it to the pitch: Combining video and movement data to enhance team sport analysis. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):13–22, 2017.

[57] M. Stein, H. Janetzko, T. Schreck, and D. A. Keim. Tackling similarity search for soccer match analysis: multimodal distance measure and interactive query definition. *IEEE Computer Graphics and Applications*, 39(5):60–71, 2019.

[58] R. Therón and L. Casares. Visual analysis of time-motion in basketball games. In *Proceedings of International Symposium on Smart Graphics*, pp. 196–207. Springer, 2010.

[59] T.-Y. Tsai, Y.-Y. Lin, H.-Y. M. Liao, and S.-K. Jeng. Recognizing offensive tactics in broadcast basketball videos via key player detection. In *IEEE International Conference on Image Processing (ICIP)*, pp. 880–884, 2017.

[60] E. Wall, S. Das, R. Chawla, B. Kalidindi, E. T. Brown, and A. Endert. Podium: Ranking data using mixed-initiative visual analytics. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):288–297, 2017.

[61] J. Wang, K. Zhao, D. Deng, A. Cao, X. Xie, Z. Zhou, H. Zhang, and Y. Wu. Tac-Simur: Tactic-based simulative visual analytics of table tennis. *IEEE Transactions on Visualization and Computer Graphics*, 26(1):407–417, 2019.

[62] K.-C. Wang and R. Zemel. Classifying NBA offensive plays using neural networks. In *Proceedings of MIT Sloan Sports Analytics Conference*, vol. 4, 2016.

[63] D. Weng, H. Zhu, J. Bao, Y. Zheng, and Y. Wu. HomeFinder Revisited: Finding ideal homes with reachability-centric multi-criteria decision making. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2018.

[64] Wikipedia. Pick and roll. `https://en.wikipedia.org/wiki/Pick_and_roll#cite_note-abrams-1`.

[65] J. Wu, Z. Guo, Z. Wang, Q. Xu, and Y. Wu. Visual analytics of multivariate event sequence data in racquet sports. In *2020 IEEE Conference on Visual Analytics Science and Technology (VAST)*, pp. 36–47, 2020.

[66] Y. Wu, D. Deng, X. Xie, M. He, J. Xu, H. Zhang, H. Zhang, and Y. Wu. OBTracker: Visual analytics of off-ball movements in basketball. *IEEE Transactions on Visualization and Computer Graphics*, 2022.

[67] Y. Wu, J. Lan, X. Shu, C. Ji, K. Zhao, J. Wang, and H. Zhang. iTTVis: Interactive visualization of table tennis data. *IEEE Transactions on Visualization and Computer Graphics*, 24(1):709–718, 2018.

[68] A. Yanez, J. Duggan, C. Hayes, M. Jilani, and M. Connolly. PandemCap: Decision support tool for epidemic management. In *2017 IEEE Workshop on Visual Analytics in Healthcare (VAHC)*, pp. 24–30, 2017.

[69] Q. Zhi, S. Lin, P. Talkad Sukumar, and R. Metoyer. GameViews: Understanding and supporting data-driven sports storytelling. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, 2019.