# SimuExplorer: Visual Exploration of Game Simulation in Table Tennis

Ji Lan, Zheng Zhou, Jiachen Wang, Hui Zhang, Xiao Xie, and Yingcai Wu

**Abstract**—We propose SimuExplorer, a visualization system to help analysts explore how player behaviors impact scoring rates in table tennis. Such analysis is indispensable for analysts and coaches, who aim to formulate training plans that can help players improve. However, it is challenging to identify the impacts of individual behaviors, as well as to understand how these impacts are generated and accumulated gradually over the course of a game. To address these challenges, we worked closely with experts who work for a top national table tennis team to design SimuExplorer. The SimuExplorer system integrates a Markov chain model to simulate individual and cumulative impacts of particular behaviors. It then provides flow and matrix views to help users visualize and interpret these impacts. We demonstrate the usefulness of the system with case studies and expert interviews. The experts think highly of the system and have obtained insights into players' behaviors using it.

**Index Terms**—Sports Visualization, Game Simulation, Model Interpretation, etc.

## 1 Introduction

Table tennis is an extremely interactive sport, in which what each player does is highly affected by the actions of his or her opponent [1]. In table tennis, players use paddles to hit a ball back and forth, using various strokes that may differ in technique, ball position, and other technical attributes. Each stroke is heavily influenced by the stroke that came before. Table tennis analysts pay close attention to how players interact through strokes, and try to determine how these interactions affect each player's ability to score, in order to help coaches formulate training plans.

Fine-grained data describing how table tennis players interact during matches is increasingly available, leading data scientists to propose a set of statistical methods and mathematical models [1], [2], [3], [4], [5], [6] to help analysts better understand player behaviors. In this domain, a "stroke state" is defined as a set of technical attributes that combine to make up a single stroke, and a "behavior" (also known as a "return") is how a player responds to his or her opponent's immediately preceding stroke. In one particularly popular model, the finite Markov chain model [5], [6], a table tennis rally is viewed as a sequence of possible stroke states, where the probability of a particular stroke state depends only on the stroke state that immediately preceded it. The model enables analysts to simulate changing the behavior of a player's opponent — in other words, how the opponent responds to a player's strokes — and explore how this impacts the player's ability to score [5], [6], [7].

However, previous studies using this model fail to explain how particular behavior adjustments actually impact scoring rates. Two studies present the impacts of different behavior adjustments simulated by the Markov chain model through static charts [5], [6]. An additional tool, Tac-Simur [7], employs a visualization system that enables users to explore the impacts of different behavior adjustments interactively. Although the explainability of the Markov chain means that the causal mechanisms that lead from behaviors to scoring rates are technically traceable, these studies do not emphasize or visualize them. This is a missed opportunity, as knowledge of these causal mechanisms allows table tennis analysts to understand when and how the behaviors of a player affect his or her scoring rate.

In this study, we present a visualization system that explicitly shows how players' stroke behaviors impact scoring rates in a table tennis simulation. Our major challenge was determining how to identify and visualize the ways in which behavioral adjustments impacted the score, both immediately and cumulatively. We were faced with three particular problems: (1) Previous studies focus on coarsely defined behaviors consisting of only one attribute per stroke [5], [6]. Table tennis analysts require a more in-depth analysis that takes both stroke technique and ball position into account over the course of two strokes. Identifying important combinations of these two attributes over two strokes in order to classify them into behaviors requires domain knowledge, time and effort. (2) Visualizing the impacts of a particular behavioral adjustment involves showing how many different entities relate to many other different entities (hereafter referred to as "many-to-many relations"). A behavior consists of two strokes — the stroke received by the player, and the stroke they make in return — and each stroke has many possible attribute values. Adjusting a behavior involves changing transfer probabilities from many attribute values of the received stroke to those of the responding stroke. A player's behavior will also influence their opponent's behavior, which further affects the scoring rate. A useful tool must effectively present these multi-step impacts and the connections between them. (3) Full table tennis rallies are made up of multiple behaviors, some of which are repeated. A useful tool must visualize not only

- *J. Lan, J. Wang, and Y. Wu are with the State Key Lab of CAD&CG, Zhejiang University. E-mail: {lanjizju, wangjiachen, ycwu}@zju.edu.cn.*
- *Z. Zhou, H. Zhang, and X. Xie are with the Department of Sport Science, Zhejiang University. E-mail: {zhouzhengzju, zhang_hui, xxie}@zju.edu.cn.*
- *Y. Wu and X. Xie are the co-corresponding authors.*

Table 1. Key Terminology.

| Term | Definition |
|---|---|
| Stroke | When a player hits the ball once with the table tennis racket. |
| Rally | A sequence of alternating strokes hit by two players. |
| Stroke Attribute | A technical element of a stroke. The two most important stroke attributes are stroke technique and ball position. |
| Stroke State | The combined values of a stroke's attributes, in this case stroke technique and ball position. |
| Behavior/Return | A player's response to his or her opponent's stroke, combined with the stroke the player is responding to. |
| Return State | The combined values of all stroke attributes involved in the two strokes that comprise a return. |
| Tactic | A tactic used by a player denotes three consecutive strokes, of which the first and third strokes are given by that player and the second stroke is given by the opposing player. |
| First Return in Tactic | The first two strokes within a tactic comprise the first return of that tactic. |
| Second Return in Tactic | The second (and final) two strokes within a tactic comprise the second return of that tactic. |
| Return Adjustment | When one player forces the other to return a specific stroke in the first return of a tactic. |
| Tactic Result | Whether a player's tactic succeeds and allows them to score. |

the immediate impacts of a behavior, but how these impacts accumulate over the course of a rally.

We worked closely with experts in table tennis to address these challenges. We identified typical behaviors. In table tennis, two overlapping behaviors (three strokes total) make up a common unit of analysis called a "tactic." We considered a rally to be a sequence of one player's tactics, and measured the impacts of that player's behavior at each tactic.

Based on these conceptual underpinnings, we modified the Markov chain model used in previous studies [5], [6] to model new behaviors. Finally, we developed a visual analytics system that provides a multi-level presentation of how different players' behavior adjustments impact their scoring rates. The system comprises four views: the player view, the adjustment view, the accumulation view, and the impact view. The player view provides an impact distribution over behaviors for different players. The adjustment view presents impacts of adjustments to various behaviors for an individual player. The accumulation view adopts flow charts to present the impact accumulation at each tactic. The impact view uses a pair of matrices to reveal the generation patterns of impacts.

The contributions of this work are as follows:

- Characterization of domain problems in interpreting impacts of behavior adjustments in table tennis.
- A visualization system that supports multi-level exploration of the impacts of behavior adjustments on scoring rates in table tennis.
- Case studies that illustrate the augmented capacity of analyzing the impacts of behavior adjustments.

## 2 Background

This section presents the domain knowledge necessary to understand the paper, along with the analysis tasks involved.

### 2.1 Domain Knowledge

Here we define key table tennis terminology used in this paper. More concise definitions of these terms are presented in Table 1.

**Stroke and rally.** A stroke occurs when a player hits the ball once with the table tennis racket. A rally is a sequence of strokes made by alternating players. When reading a particular rally's stroke sequence, strokes with odd sequential numbers were made by Player A and strokes with even sequential numbers were made by Player B. A full table tennis match comprises nearly a hundred rallies.

**Stroke attribute.** A single stroke has many technical attributes. According to table tennis analysts, the two most important attributes are stroke technique (how the player hits the ball) and ball position (where the ball impacts the table after being struck). In this study, we use these two attributes to describe each stroke.

**Stroke state.** The combined values of a stroke's two attributes — in this case, stroke technique and ball position. For instance, $(topspin, long\,forehand)$ is a stroke state in which the stroke technique is *topspin* and the ball position is *long forehand*.

**Behavior/Return.** With the exception of the initial serve, each stroke made in a table tennis rally follows a previous stroke made by an opponent. A player observes their opponent's stroke and considers how to respond to it with their own stroke. To table tennis experts, this set of two strokes — the prior stroke and the response — is called a "return." In this paper, because it is the unit of analysis our system works with, we also refer to it as a "behavior." In table tennis analysis, a return also contains information about the prior stroke. A return by Player B is formally denoted $(stroke_{i,p1}, stroke_{i+1,p2})$.

**Tactic.** Table tennis players think several steps ahead, using their own returns to try to force particular returns from their opponents in order to set themselves up for a specific return of their own. In table tennis analysis, a "tactic" consists of three total strokes — two by the player at hand and the middle one by their opponent. Formally, $(stroke_{i,p1}, stroke_{i+1,p2}, stroke_{i+2,p1})$ is a tactic by Player A. The first two strokes of a tactic comprise that tactic's "first return" $(stroke_{i,p1}, stroke_{i+1,p2})$. The second (and last) two strokes in a tactic comprise its "second return" $(stroke_{i+1,p2}, stroke_{i+2,p1})$.

**Return adjustment.** Within a tactic, Player A forces Player B to return a specific "setup" stroke in the first return so that they can return a planned stroke in the second return and increase their chances of scoring. We call this a "return adjustment."

**Tactic result.** A tactic by Player A has two possible results: Player A either does or does not score on the third stroke. Analysts aim to investigate how adjustments to the first return of a tactic affect the result of the tactic overall.

### 2.2 Analysis Tasks

We collaborated with two experts (hereafter referred to as "expert A" and "expert B") for six months to develop Simu-Explorer. Expert A is an authority on technique and tactic analysis of table tennis players. Expert B is a postdoctoral researcher majoring in sports science, as well as a table tennis athlete and experienced analyst. Both experts work for a top national table tennis team.

The experts hope to analyze table tennis data and interpret the results. To identify problems with their existing practices and opportunities for improvement, we conducted informal interviews, used their analysis software, and observed them at work. The experts used a Markov chain model to analyze data, often using software to adjust players' strokes in the model and examine how this impacted scoring rates. We noted that this process was often tedious, and might involve changing the behavior of every single player in order to determine how to positively impact the scoring rate (T1), or changing every single return of a player to find out which tweaks were effective (T2). Moreover, experts often found it difficult to interpret why changing a particular stroke attribute affected the scoring rate, as the software they used outputted results without providing explanations. They hoped that we could provide a visualization tool that would help them understand the relationship between a behavior change and its impacts (T3–4). We worked closely with the experts to develop such a tool. The main stages of our collaboration are as follows.

**Applying a prototype system (two months).** Experts were unfamiliar with the intermediate results of the Markov chain model, which help explain how player behaviors translate into scores. We developed a prototype system to help them explore the distribution of these intermediate results, using a Sankey diagram to present how stroke states lead to other stroke states.

They quickly understood the possibilities, pointing out that we could break up a rally into a sequence of tactics and examine how impacts of a successful return adjustment accumulate at each tactic (T3).

We further characterized the process through which an adjustment affects the scoring rate within a tactic. According to experts, an adjustment in the first return of a tactic increases the chance that an opponent uses a specific kind of stroke, and decreases the chance that the opponent uses other kinds of strokes. This adjustment then changes the probabilities of different kinds of strokes in the second return, and these changed probabilities change the scoring rates by the third stroke of the tactic. Overall, tracing the route from an adjustment to its impact involves tracking the adjusted many-to-many state probabilities in the first return, the affected many-to-many state probabilities in the second return, and the changed final scoring rate after the third stroke. The chance of scoring at a particular tactic can also be influenced by return adjustments in previous tactics. The experts need the visualization system to support visual exploration of how changed state probabilities coalesce into an impact on the chance of scoring at a tactic, and how the impact is indirectly influenced by previous tactics (T4).

We found that the prototype system had too much visual clutter, as the Sankey diagram cannot present relations among too many groups, and showed all the information instead of just necessary information.

**Characterization of domain problems (one month).** We discussed the Markov chain model with experts, and decided which intermediate model results were the most important. In this process, we referred to previous studies [5], [6] that have used the Markov chain model to simulate the impacts of adjusting players' behaviors.

**Iteration of system design (two months).** We iterated and refined our design based on multiple discussions with our expert collaborators, focusing on clarifying the information pipeline and revising and synthesizing analysis tasks.

**Development of the system (one month).** After the experts confirmed the final set of analysis tasks, we began developing the alpha version of the system. We then invited the experts to use the system and revised it into a beta version according to their feedback on functionality and usability. We continued polishing this version of the system.

After these iterative discussions and revisions, the experts settled on the following analysis tasks:

- T1 ***What are the impacts brought by return adjustments of different players?*** Players have individual styles of playing, and the impacts of returns differ by player. The experts hope to browse different players' impact distributions over return adjustments.
- T2 ***What are the impacts brought by different return adjustments of a player?*** Adjustments to different returns exert varying impacts for a player. The experts need to detect the impact patterns of different return adjustments.
- T3 ***How do the impacts brought by a return adjustment accumulate over tactics in a rally?*** A rally in table tennis contains a sequence of tactics. The experts hope to explore how the impacts of a return adjustment in sequential tactics accumulate in a rally.
- T4 ***How does a return adjustment influence the scoring chances at a tactic?*** The experts hope to examine how the impacts are generated in a tactic and how many impacts in a tactic are due to previous tactics.

## 3 Related Work

This section presents analytical methods used with table tennis data, visualizations used with sports data, and visualization techniques for displaying event sequences and state transitions.

### 3.1 Analytical Methods for Table Tennis

A common approach to evaluating table tennis players' performances is to describe statistics associated with performance indicators. Existing studies analyze shot characteristics [2], game structure indicators [3], and skill evaluation indicators [4]. These studies investigate the statistics drawn from game observations. However, such statistics ignore dynamic interactions between players and do not allow for systematic analysis [1]. The Markov chain model can be used to simulate full table tennis rallies and to identify behaviors that strongly impact a player's scoring rate. This model considers the context linking individual observations. On this basis, Pfeiffer et al. [5] developed four different state-transition models to describe the impacts of behaviors on a table tennis match. Wenninger et al. [6] used a similar model to reveal the scoring rate factors. Tac-Simur [7], which is also based on the Markov chain model, proposes the first visualization system that allows for interactive adjustment and impact examination. However, this tool does not help table tennis experts understand how adjustments impact the score, or how they build on each other.

This work goes beyond these previous studies by introducing a method for exploring how the estimated impacts of behavior adjustments on the score occur and accumulate over the course of a rally, through a modified Markov chain model and a set of coordinated visualizations.

### 3.2 Sports Visualization

A number of previous studies have tackled the problem of sports visualization. We classify these into three groups based on each study's design goals.

**Seeking spatial and temporal patterns.** Most sports visualization methods aim to detect spatial patterns in sports data. Counterpoints [8] and GameFlow [9] visually represent spatial patterns of basketball players' statistics. Baseball4D [10] and SnapShot [11] present the spatial patterns of balls and players in baseball and shot lengths in ice hockey, respectively. StatCast Dashboard [12] explores the spatial distributions of various kinds of data in baseball. PassVizor [13] presents the spatial patterns of passing behaviors in soccer. Previous studies have also explored trajectories in sports in order to obtain the attack patterns of teams [14], detect anomalous events [15], and examine the movements of balls [16], [17], [18]. Temporal patterns in sports data are also worth exploring. BKViz [19] reveals the patterns of temporal observations, such as the play types and point outcomes, of a basketball team in a match and throughout a season, respectively. TenniVis [20], CourtTime [21] and iTTVis [22] present the temporal patterns of varying scores and rally lengths in tennis and table tennis, respectively. ForVizor [23] visualizes spatiotemporal patterns of formations in soccer, and TideGrapher [24] visualizes the similar patterns for rugby football. These studies have utilized high-bandwidth visual channels to deliver valuable information to sports experts.

**Revealing patterns of relations and structures.** Visualizations that reveal relationships within sports data have also been proposed. Several studies have investigated patterns in team rankings — including A Table [25], which reveals the evolving rankings of soccer teams — while many studies also explore the tree structure of matches or tournaments. Tan et al. [26] visualized the tree structure of a tournament to provide an understandable representation of the process and allow nonlinear predictions. TennisViewer [27] provides clear tree structures of tennis match data. Two studies aim to display many-to-many relations within sports data. SoccerStories [14] and iTTVis [22] use matrix views to visualize passing rates among players in a soccer match and the correlations among attribute values over strokes in table tennis, respectively. These studies help users clarify and analyze the complex relationships within and between sports datasets.

**Glyph-based and annotation-aided analysis.** Many sports visualization tools seek to provide intuitive analysis through the use of representative glyphs, or by directly annotating video. MatchPad [28], iTTVis [22], Tac-Simur [7], Tac-Miner [29], and RallyComparator [30] each uses a set of glyphs that imitate the actions of players, to help domain analysts grasp the events in a rugby game or understand table tennis stroke attributes at a glance. Similarly, TacticFlow [31] represents tactic variation patterns through tailored glyphs, as does TenniVis [20] for game elements such as scoring and serve information. Data videos [32] and visualization-augmented videos [33] are widely employed to communicate insights engagingly. Studies have been proposed to integrate visualizations into original sports videos. Parry et al. [34] introduced a video storyboard that can be used to visually depict and annotate events in snooker videos. VisCommentator [35] proposes a design space and a prototyping tool for augmenting sport videos with visualizations. Director's Cut [36] and Bring it to the Pitch [37] use computer vision techniques to extract trajectory data from soccer match videos and add visualizations that represent important statistics, such as distance and area. These studies integrate real, often visible elements of sports into the data analysis process, making it more intuitive.

A recent survey [38] alternately classified sports data visualization efforts into feature presentation, comparison, and prediction tasks. Our work is unique in that it aims to use interactive visualizations to explain impacts within a simulation model, an area not investigated by previous sports visualization studies.

### 3.3 Visualizing Many-to-Many Relations

Studies aiming to visualize many-to-many relations mainly focus on the visual analysis of state-transition graphs, geographically-embedded flows, and event sequences. We present relevant studies in these areas.

Most previous studies focus on visualizing state transitions among many-state values have used node-link diagrams to present transition relations. For instance, PivotGraph [39] places different types of nodes in a grid and uses links among the nodes to represent relations. However, node-link diagrams are not easily scalable. Pretorius and Van Wijk [40], [41] employed dense visualization views to present overall trends within thousands of transition relations, while Van Ham et al. [42] clustered a large diagram to reduce the information. However, examining a specific relation within a dense or aggregation view is impossible. For this study, where table tennis experts need to examine a specific relation among the many relations, these visualization techniques cannot be applied.

The visualization of geographically-embedded flows aims to show movements between and among many geographic locations. Previous studies have employed flow maps [43] to represent many-to-many relations among locations. A flow map places locations in a map according to their geographical coordinates, and connects these locations with links whose width encodes the traffic volume. The flow map is intuitive but also has limited scalability. An origin-destination (OD) matrix [44] may also be proposed to visualize directed movements from many origins to many destinations. The matrix view possesses high information density and clutter-free features, and is more scalable than a flow map for visualizing many-to-many relations. Because an OD matrix does not contain geographical information, an OD map [45] has been proposed that includes this, while another study [46] proposed the use of leader lines to link an OD matrix and its corresponding map. These visualization techniques are closely tied to geographical locations, and do not consider the temporal component of many-to-many relations. The data in our study do not contain geographical information, and we must include temporal analysis of these many-to-many relations, so these strategies are not appropriate for our task.

Visualizing multiple event sequences involves showing the transition relations among many event types in time. Moreover, the relations within event sequences have a temporal order. Many studies employ flow charts [47], [48] to visualize sequential many-to-many relations. Although flow charts are intuitive, they can also run into scalability problems. In our study, there are more than two hundred possible relations among twenty attributes. A flow chart used to visualize relationships on this scale will be visually cluttered. MatrixWave [49] and iTTVis [22] employ a matrix flow to visualize sequential many-to-many relations [22], [49]. The matrix view is more scalable for presenting relationships, and its tailored layout allows for the display of relations within different time steps. But the layouts of the sequential matrices in MatrixWave [49] and iTTVis [22] change the direction of a matrix flow, which can hinder understanding of how impacts build on each other during a rally. Moreover, these two layouts do not save space for displaying two matrices in a rectangular view. In this study, we employ matrix views to visualize the multi-step, many-to-many relationships between player behavior attributes and scoring impacts, and propose a new layout to better present the sequential impacts.

# 4 Design and Development of SimuExplorer

SimuExplorer is a web application with three components: one each for data preprocessing, data analysis, and visualization. The data preprocessing component extracts typical returns of players (Fig. 1A) from the CSV data tables that record matches. The data analysis component builds a modified Markov chain model (discussed in Section 4.2) for each player (Fig. 1B). For each return adjustment of each player, the model estimates the overall impact on the scoring rate (Fig. 1B). The model also outputs intermediate results that explain the generation process of impacts (Fig. 1B). The visualization component (Fig. 1C) uses Vue.js to interactively visualize the multi-level impacts and intermediate results received from the model. We provide a demo of SimuExplorer on *https://simuexplorer.github.io/*.

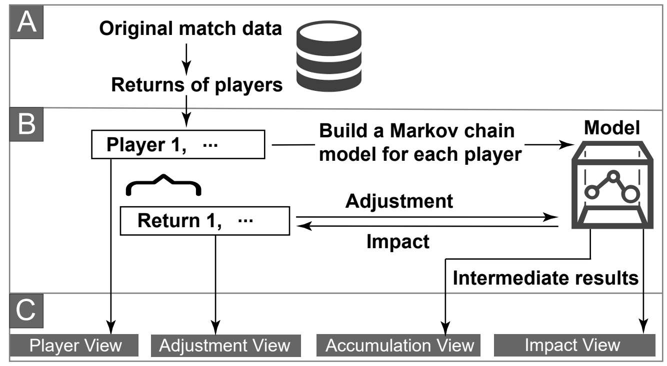


Fig. 1. System Overview. (A) Data processing component. (B) Data analysis component. For each player, this component builds a Markov chain model. The impact of each return adjustment of the player is estimated by the model. (C) Visualization component. This component visualizes the multi-level impacts output by the model.

## 4.1 Data

We collected data from 306 table tennis matches among 21 top players (10 males and 11 females) from 2005 to 2012. The data was collected by manually coding videos. We have tried using machine learning models to automatically extract structured data from video. However, low video quality means that the accuracy of automatically extracted stroke features, such as stroke technique, is relatively low (around 75%). Each match was collected as a CSV file, which contained hundreds of rows representing the hundreds of strokes in the match. Each row records different features of a stroke as follows.

- **Rally ID** denotes the rally the stroke belongs to.
- **Stroke ID** is the sequence number of the stroke in the rally.
- **Stroke player** denotes the player who gave the stroke.
- **Stroke technique** denotes the technique the player uses to give the stroke. There are 14 techniques.
- **Ball position** denotes the drop point of the stroke. There are ten drop points.

## 4.2 Simulative Model

Lames [6] first proposed simulating a table tennis rally through a Markov chain model. The model views a rally as a sequence of possible stroke states, and simulates the impacts of strokes on players' scoring rates. Here we go through how Lames' model [6] has been applied in previous studies and explain how the model simulates a rally. Then we propose a modified model and explain how intermediate results of the model measure returns and impacts of return adjustments.

### 4.2.1 Lames' Model

Previous studies [5], [6] use a stroke attribute, such as stroke technique, as the transition state of the Markov chain. In this manner, a rally in table tennis is regarded as a sequence of possible stroke attribute values on $stroke_1$, $stroke_2$, ... $stroke_n$. Formally, the states are $value_{1,p1}$, $value_{1,p2}$, $\ldots$, $value_{n,p1}$, $value_{n,p2}$ , $score_{p1}$, and $score_{p2}$. These states are all attribute values of the selected stroke attribute (e.g., stroke technique) of two players (P1 and P2). $score_{p1}$ and

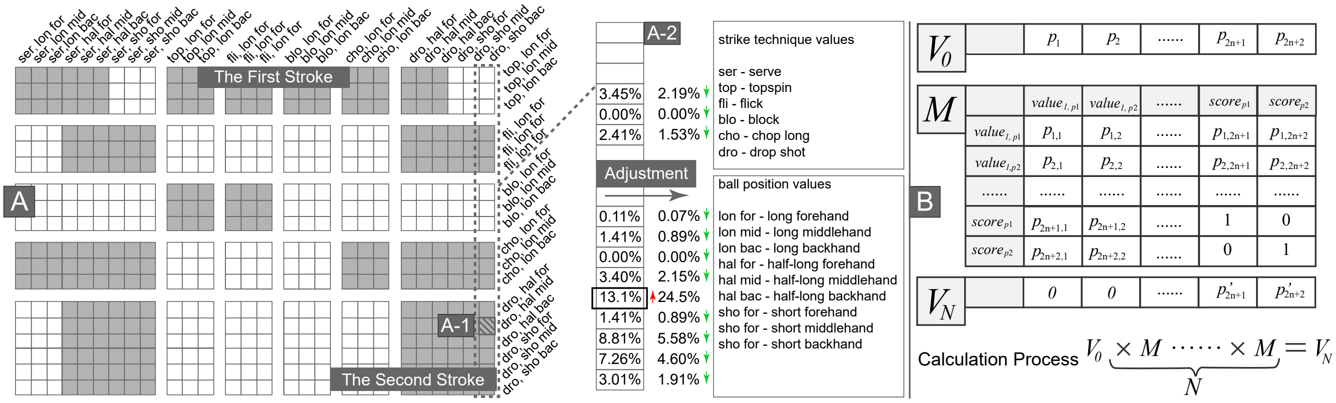


Fig. 2. (A) We array 243 two-stroke returns as gray entries in a matrix. The column headers of the matrix are value combinations of stroke technique and ball position at the first stroke of the return, and row headers are those at the second stroke. (A-1) is a highlighted return. (A-2) gives an example of adjusting the return. (B) Illustration of the Markov chain model, where $M$ is the transition matrix and $V_0$ and $V_N$ are the state probability vectors.

$score_{p2}$ indicate that the player has scored at that stroke. These two states are the absorbing states and will not transform into other states.

As shown in Fig. 2B, $M$ is the *empirical transition matrix* and $V_0$ is the *empirical initial probability vector.* The row and column headers of the matrix are the states, and each matrix entry presents the transition probability from the row header to the column header. Each element in vector $V_0$ presents the initial probability of the corresponding state. The transition matrix and initial probability vector are estimated from the analyzed matches. We calculate the ratio between the number of times that state A transforms into state B and the number of times that state A transforms into all states in collected matches, and use this ratio as the probability in the matrix entry with row header A and column header B. We calculate the ratio between the number of times that state A appears in a rally's first stroke and the number of times that all states appear in the first stroke. We use this ratio as the probability in the initial probability vector $V_0$. During computation, the initial probability vector is repeatedly multiplied by the transition matrix until the probabilities of all states except absorbing states approach zero (to five decimal points). The probabilities of $score_{p1}$ and $score_{p2}$ in the absorbing vector $V_N$ are regarded as the scoring rates of the two players. The transition matrix $M$ measures player returns. Adjusting a player's returns is achieved by tuning the empirical transition matrix $M$. Changes in final scoring rates then quantify the impacts. The specifics of these adjustments will be introduced in the next section.

However, a player's stroke is complex, and many key stroke features involve combinations of two attributes, which previous studies cannot model. Moreover, a player's return is also dependent on the opponent's prior return within a tactic. Merely modeling transitions among strokes loses helpful information about tactical correlations between returns, thus inadequately quantifying the impacts of return adjustments.

### 4.2.2 Modified Model

We modified the model to consider two attributes and two strokes as a single state in order to preserve the complex stroke features and tactical correlations among returns.

**Measuring return.** In collaboration with the experts, we decided to use the stroke technique and ball position to describe each stroke, and to use two consecutive strokes (one return) as a single "return state." After discussions and explorations of the collected data, we selected 243 typical returns (Fig. 2A).

We used player returns as transition states in order to preserve tactical correlations between returns. In this manner, a rally is regarded as a sequence of possible return states on $stroke_{1,2}$, $stroke_{2,3}$, ... $stroke_{n-1,n}$. The transition matrix $M$ and state probability vector $V$ are estimated similar to Lames' model.

**Making adjustments.** We break down a rally into the tactics of the player who gives the first stroke, i.e., $stroke_{1-3}$, $stroke_{3-5}$, ... $stroke_{2k+1-2k+3}$. Each tactic is further divided into two returns, i.e., $(stroke_{1-2}, stroke_{2-3})$, $(stroke_{3-4}, stroke_{4-5})$, ... $(stroke_{2k+1-2k+2}, stroke_{2k+2-2k+3})$. As introduced in Section 2.1, a player will adjust the return of her/his opponent during the first return of the tactic in order to obtain a higher chance of scoring after the third stroke. We hence allow adjustment of the first return in a tactic (e.g., the $stroke_{1-2}$ of the tactic $stroke_{1-3}$) in the model and explore how this impacts the second return (e.g., the $stroke_{2-3}$ of the tactic $stroke_{1-3}$), and finally the scoring rate after the third stroke.

As shown in Fig. 3B, adjustments $\Delta V_0$ and $\Delta V_2$ are applied to the probability vectors at returns comprising Strokes 1–2 and Strokes 3–4, respectively. The probability vectors at returns comprising Strokes 2–3 and Strokes 4–5 receive a boost of $\Delta V_0 M$ and $\Delta V_0 M^3 + \Delta V_2 M$ (Fig. 3A), respectively. The scoring rates at Strokes 3 and 5 are subsequently influenced (Fig. 3C). In this manner, we can determine overall how the return adjustments of player A during the first returns of her or his tactics change the probabilities of return states at the second returns, and finally change her or his chance of scoring by the end of the tactics. In particular, for the variations in probability vectors at the return comprising Strokes 4–5, $\Delta V_2 M$ is caused by the return adjustment at the return comprising Strokes 3–4, and $\Delta V_0 M^3$ is caused by adjustments at previous tactics. Our visualization system distinguish these two impacts.

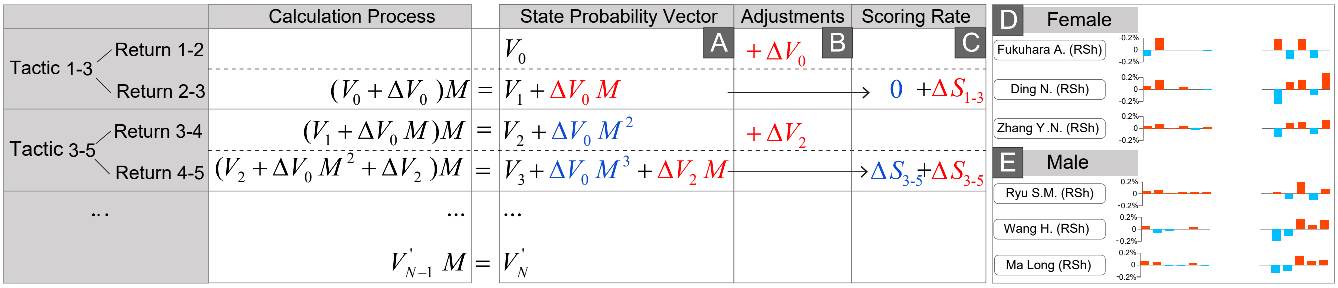


Fig. 3. Illustration of the adjustments and impacts during the computation process of the modified model. Items in red are caused by adjustments at the current tactic, and items in blue are caused by adjustments at previous tactics. (A) presents the changed state probability vector at each step in the Markov chain process. (B) presents the adjustment made at each tactic. (C) presents the changed scoring rate at each step in the Markov chain process. (D) and (E) present the distributions of impacts over return adjustments of three female and three male players for the case study discussed in Section 5.1.1.

We calculated the magnitude and compensation of adjustments similarly to previous studies [5]. For instance, assume the experts want to see what happens if player A forces the opponent to use $(drop, half\,middlehand)$ stroke state more frequently after he or she uses $(drop, short\,backhand)$ stroke state (Fig. 2A-1). We firstly increase the probability of this stroke state and reduce probabilities of other stroke states in the same column proportionately, to guarantee that the total probability of the column is constant (Fig. 2A-2). According to Pfeiffer et al. [5], the function for deflection is as follows (All probabilities in the column are normalized through division by the total probability of the column):

$$\delta P_x = C + B * 4 * P_x * (1 - P_x)$$

where $P_x$ is the normalized probability of the state; $\delta P_x$ is the change of the probability; $C$ is a constant that describes the deflection in the border probabilities; $B$ is a constant that describes the maximum value of the relative magnitude of deflection; and 4 is a normalization factor that allows the constant $B$ to be equal to the maximum value of deflection. In this work, the constant C = 0.05 and B = 0.25, determined on the basis of the previous work [5] and discussions with the experts. According to Pfeiffer et al. [5], the compensation function is

$$\delta P_{yi} = -(P_{yi}/(1 - P_x)) * \delta P_x$$

where $P_{yi}$ is the normalized probability of other states in the column and $\delta P_{yi}$ is the change in probability. Fig. 2A-2 shows an example for the adjustments.

## 4.3 Visualization Design

We propose a visualization system to realize the analysis tasks proposed in Section 2.2. The system consists of four views: the player view, the adjustment view, the accumulation view, and the impact view. The workflow and cross-view interactions are as follows.

- *Click a player in the player view, and examine corresponding impacts in the adjustment view.* Bar charts in the player view represent the impact distribution over different returns for a chosen player (Fig. 4A, T1). When a user is interested in the impact distribution of a particular player, she/he can click and select the player in the player view. The adjustment view will then display the impacts of adjusting different returns for the player (Fig. 4B, T2).
- *Click an impact in the adjustment view and examine its accumulation in the accumulation view.* The user can further select the impact of an interesting return adjustment by clicking it in the adjustment view. The accumulation view provides visual tracking of the accumulation of impacts over tactics (Fig. 4C, T3).
- *Click a tactic in the accumulation view and examine how the impact is generated at this tactic in the impact view.* When the user clicks a tactic in the accumulation view, the impact view presents the generation process of the impact at that tactic (Fig. 4D, T4).

The four views provide a multi-level presentation of the impacts caused by adjusting players' returns. **Green** and **purple** represent the two players, and **orange** and **blue** represent the increase and decrease in the scoring rate and probability over the whole system.

### 4.3.1 Player View

The player view (Fig. 4A) provides a browsable overview of the impacts of return adjustments for different players (T1).

This view contains a table that lists all the top table tennis players. A bar chart (Fig. 4A-4) in each row in the table presents the impact distribution of return adjustments for a player. Because stroke technique is the most important attribute, in the left column, we group typical returns according to six stroke technique values based on the first stroke in the return, obtaining six groups. In the right column, we group typical returns according to five stroke technique values based on the second stroke in the return, obtaining five groups. We encode the average impact of adjusting returns in each group as a bar in the bar chart. The orange bars denote return adjustments with positive impacts, while the blue bars denote return adjustments with negative impacts. The axes are scaled according to the largest impact. With this view, users can quickly detect whether a player can change their scoring rate by adjusting a group of returns involving specific stroke technique values. Two switch buttons (Fig. 4A-1 and 4A-2) allow switching between serve and other rallies and between male and female players.

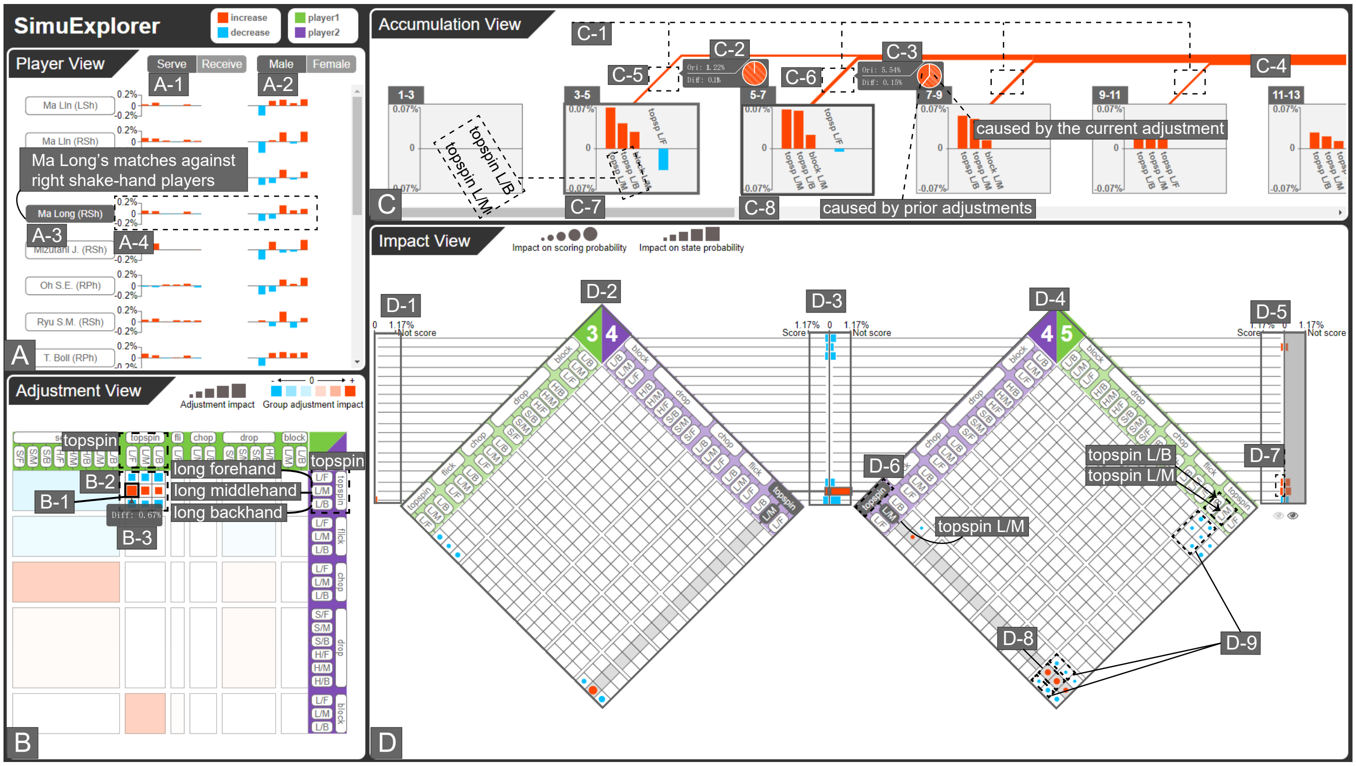


Fig. 4. System Interface. The interface comprises four well-coordinated views: the player view (A), the adjustment view (B), the accumulation view (C), and the impact view (D) (including a pair of coordinated matrices (D-2) and (D-4)).

#### *4.3.2 Adjustment View*

After a user selects a player for analysis, the adjustment view (Fig. 4B) provides an overview of the impacts of adjusting the player's different returns (T2).

This view contains a matrix similar to that in Fig. 2A. Each entry in the matrix (Fig. 4B-1) represents one return of Player A. The column headers denote the combinations of stroke technique and ball position values at the first stroke of the return (given by Player A), and row headers denote those at the second stroke (given by opponents of Player A). The area of the rectangle in an entry encodes the impact of adjusting the corresponding return. (To reiterate, adjusting a return means successfully influencing the opponents of Player A to use a specific type of second stroke more frequently after Player A uses another specific type of stroke.) The matrix hides any rows and columns with small values in order to emphasize key return adjustments. To ease navigation of this large amount of returns, we group columns and rows according to the stroke technique of the first and second stroke in the return, respectively. Initially, detailed information for each entry is not visible. Instead, group blocks are displayed, where the lightness of a block encodes the average impacts of returns within that block. If a user is interested in a block of returns, she/he can click the block and examine the detailed impacts of each return within the block (Fig. 4B-2). When a user hovers over an entry, the size of the impact will be displayed (Fig. 4B-3). Users can also click an entry to select particular a return adjustment of Player A, to explore how the impacts of this adjustment are generated and accumulated.

#### *4.3.3 Accumulation View*

After a user selects a return adjustment of Player A, the accumulation view (Fig. 4C) enables visual tracking of how the impacts of this return adjustment accumulate in each tactic (T3).

This view contains a flow chart that represents the variation in Player A's scoring rate after he or she adjusts a return during his or her tactics (Fig. 3C). Flow charts are widely used to visualize temporal variations in variables [50], [51], [52]. In the accumulation view, the flow chart intuitively illustrates the process through which impacts accumulate gradually over the course of different tactics and finally aggregate into the total impact. In the accumulation view, the width of the main flow (Fig. 4C-4) encodes the accumulated changed scoring rate, and the width of branches (Fig. 4C-1) encodes the variation of the scoring rate at each tactic. When users hover on each branch, the original scoring rate (above) and changed scoring rate (below) are displayed in a detail view (Fig. 4C-3). For the changed scoring rate, a pie chart displays how much of this rate change was caused by the adjustment at the current tactic versus how much was caused by adjustments at previous tactics.

The bar chart (Fig. 4C-7) of each tactic shows six stroke states, arranged based on their impact on scoring probabilities when they are at that tactic's last stroke. The first three states displayed are those that do the most to increase scoring probabilities. The final three displayed are those that do the most to decrease those probabilities. The height of the orange bar encodes the increase in scoring probability, while the height of the blue bar encodes the decrease. This helps users to see how adjustments impact a tactic's last stroke and the corresponding scoring probabilities.

#### 4.3.4 Impact View

After a user selects a tactic, the impact view (Fig. 4D) presents how the impact is generated through the two returns within the tactic. The impact view must display the adjusted many-to-many probabilities of states in the first return, the changed probabilities of states in the second return, and the scoring rate after the third stroke (T4).

**Justification:** There exist more than two hundred return states, with each state involving a specific first stroke and second stroke (Fig. 2). Displaying probabilities of return states is a many-to-many relation visualization problem.

***Flow or matrix.*** As discussed in Section 3.3, existing visualization studies employ node-link diagrams, flow charts, and matrix views [53] to display many-to-many relations. We initially employed flow charts to present the changed probabilities, thinking they would be intuitive to understand. However, these ended up visually cluttered due to the many probabilities involved. We thus switched to a clutter-free matrix view to help users browse how probabilities were changed by the adjustments.

***Layout.*** Two matrices display changing probabilities over two returns. As discussed in Section 3.3, the layouts used by MatrixWave [49] and iTTVis [22] were not suitable for this project. Instead, we propose our current design, which includes clutter-free matrix views and straight leader lines with interactions that help users understand and interpret how impacts compile within a tactic. The details of the design are introduced as follows.

**Description:** The impact view uses a pair of matrices to represent the changed probabilities of the states at the two returns of a tactic, thus explaining the varying scoring rate after the third stroke of this tactic. The first matrix (Fig. 4D-2) presents how adjustments are made at the first return of the tactic. The second matrix (Fig. 4D-4) displays how the adjustment changes the probabilities of states at the second return and further changes the scoring rate after the third stroke. The three columns of *stroke bars* (Fig. 4D-1, 4D-3, and 4D-5) from left to right represent the changed probabilities of stroke states (different stroke technique and ball position values) at the three strokes in a tactic. These probabilities in each bar column are derived by aggregating the probabilities in corresponding columns in the prior matrix. The detailed encodings are as follows.

***Bars for strokes.*** Three columns of bars (Fig. 4D-1, 4D-3, and 4D-5) present the changed probabilities of stroke states at the first, second, and third strokes. The height of each bar encodes the changed probability of each stroke state. The bar color indicates whether the probability increases (orange) or decreases (blue). The probability of each stroke state is divided into two parts. The first part is the changed scoring probability after the stroke, i.e., the change in the probability that the player gives a stroke with this state and scores directly. This is encoded by bars on the left. (This part of the first stroke is not shown, because it is not considered in the current tactic). The second part is the change in the probability that the player gives a stroke with this state and does not score. This part is encoded by bars on the right. When users hover on each part of a bar, the corresponding rows and columns in the two matrices are highlighted (Fig. 4D). The original and changed probabilities are also displayed in the detail view (Fig. 5C). A pie chart displays how much of this probability change was caused by the adjustment at the current tactic versus how much was caused by adjustments at previous tactics.

***Matrices for returns.*** A pair of matrices presents the changed probabilities of all states that make up the two returns within a tactic. The first matrix (Fig. 4D-2) presents the adjustments (Fig. 3B) to the first return in the tactic. The second matrix (Fig. 4D-4) presents the impacts on the second return of the tactic (Fig. 3A). Similar to the layout of return states in Fig. 2A, the changed probabilities of return states are arranged as a matrix whose column and row headers represent the stroke state at the first and second of the two strokes, respectively. The headers are linked by bars representing aggregate changed probabilities. As the impacts pass through the matrix, the changed probabilities of stroke states branch into the changed probabilities of return states, and then merge into the changed probabilities of the next stroke states. With this detailed display of changed probabilities of return states, experts can better understand how state probabilities of a return are adjusted, how adjustments to previous returns change the state probabilities of following returns, and how these changed probabilities aggregate into changed scoring rates. When the serve tactic (i.e., the tactic comprising Strokes 1–3) is presented in the impact view, the columns of the first matrix are altered because the first stroke in the serve tactic must be a serve stroke.

In each matrix entry, the area of the circle encodes the changed probability as it transforms from the column header to the row header. The circle color indicates whether the probability increases (orange) or decreases (blue). A switch button is placed at the bottom of the bars for the third stroke. The right part of the button (corresponding to the right part of the bars) is enabled by default. Circles in entries of the second matrix represent the changed probabilities from the column headers to the row headers. When users switch to the left part of the button (which corresponds to the left part of the bars), the circles in the entries of the second matrix are transformed into rectangles. The areas of the rectangles represent the changed scoring probabilities at the third stroke due to the changed probabilities of the first and second. When users hover on an entry, the corresponding row and column are highlighted, and a detail view (Fig. 5B) is displayed.

## 5 Evaluation

We invited two collaborating experts (hereafter referred to as “expert A” and “expert B”) introduced in Section 2.2 and two new experts (hereafter referred to as “expert C” and “expert D”) to evaluate the system. The two new experts were not involved in developing the system. Expert C is a Ph.D. student in sports science and expert D is a master’s student. Both of them are senior analysts of table tennis data.

We gave a tutorial on how to explore the impacts on the system to the experts. (1) We introduced what problems and analysis tasks this study targets to the new experts (10 minutes). (2) We described the visualization designs of SimuExplorer and demonstrated how to use the system to complete the analysis tasks (15 minutes). (3) The experts explored the system by themselves. We answered their questions during the exploration (20 minutes).

We let the experts use the system to analyze the data (introduced in Section 4.1) for two weeks after the experts could use the system correctly. During this time, we answered their questions at any time online. Afterward, the experts demonstrated to us how they found insightful patterns using the system. We summarized the exploration processes that experts A and B identified two patterns as two case studies (Sections 5.1.1 and 5.1.2). We also summarized the exploration processes that experts C and D identified an insightful pattern as the third case study (Section 5.1.3). These case studies were conducted on Google Chrome on a PC (equipped with a $1920 \times 1080$ display). Besides the insights in the case studies, we also presented other insights obtained by the experts (Section 5.2). Finally, we asked them open-ended questions about the usefulness and usability of our system and collected their feedback (15 minutes, Section 5.3). The detailed open-ended questions can be referred to in the supplementary materials.

## 5.1 Case Studies

We present three case studies to evaluate the system.

### *5.1.1 Impact Distributions of Different Players*

In this case study, the expert examined different players' impact distributions over return adjustments in the player view. He then selected three female and male players due to their distinct features. Fig. 3D illustrates the impact distributions of three female players, Fukuhara, Ding, and Zhang. (For each player, the impacts were estimated according to matches between the player and female right shake-hand players.) The expert found that the positive impacts of adjusting Ding and Fukuhara's returns were larger than those of adjusting Zhang's returns. He interpreted this result that fewer improvements could be made in Zhang's returns than those of Ding and Fukuhara. Zhang performed better than the other two players. Fig. 3E illustrates the impact distributions of three male players, Ryu, Wang, and Ma. The expert found that the positive impacts of adjusting Ma and Wang's returns were smaller than those of adjusting Ryu's returns. The expert interpreted the result that Ma and Wang performed better than Ryu because fewer improvements could be made in their returns.

This case study demonstrates that an expert could browse and evaluate different players' impact distributions over return adjustments using the player view (Fig. 3D and 3E). The software used by the experts (introduced in Section 2.2) cannot provide such an overview.

### *5.1.2 Good Performance in Receiving Topspin to Middlehand*

This case study analyzes matches between Ma Long and right shake-hand players. Ma Long is commonly regarded as one of the greatest table tennis players and has won the Olympic games twice. The expert selected Ma Long and right shake-hand players in the player view (Fig. 4A-3) and browsed the adjustment view afterwards. He focused on the return adjustments (Fig. 4B-2) that let the opponents use strokes with *topspin* (a stroke technique value) more frequently after Ma Long's strokes with *topspin*. Then he found that the scoring rate always increased when the opponents' strokes were to *long middlehand* (a ball position value) as the rectangle colors are orange. And the scoring rate always decreased when the strokes were to *long backhand* and *long forehand* (two ball position values) as the rectangle colors are blue. The expert commented that when a male player let the opponents use strokes with *topspin* more frequently he would decrease his scoring rate. It is because he would be less likely to employ strokes with *topspin* at the third stroke and score in that case. Hence, it is against common knowledge that the scoring rate increased. The expert hoped to determine the reasons. He clicked the biggest one of the three orange rectangles. (Fig. 4B-1).

The accumulation view displays how the impact accumulated in a rally (Fig. 4C). The expert found the impact accumulated at each tactic evenly (Fig. 4C-1). He browsed the bar chart of the tactic comprising Strokes 3–5 (Fig. 4C-5) because the scoring probability began to increase in the tactic. In this tactic, the scoring probability increased the most when the third stroke of the tactic was with *topspin* to *long middlehand* (hereafter referred to as "*topspin L/M*" ). It increased the second most when the third stroke was with *topspin* to *long backhand* (hereafter referred to as "*topspin L/B*" ). The expert clicked this tactic (Fig. 4C-7) to examine how these increments were generated.

In the impact view (Fig. 4D), he found that the adjustment increased the probability of opponents' stroke with *topspin L/M* (Fig. 4D-6). This increment then increased the probabilities of Ma's next stroke with *topspin L/M* and *topspin L/B* (Fig. 4D-8), which further increased the chances that Ma scored after the tactic (Fig. 4D-8). Moreover, the decrease in the probabilities that Ma gave back with *topspin L/M* and *topspin L/B* (Fig. 4D-9) did not exceed the increase. Therefore, Ma had more chances to score after this tactic (Fig. 4D-7).

The expert browsed the accumulation view again and found that the increment in the scoring rate at the tactic comprising Strokes 5–7 (Fig. 4C-6) was larger than that of the tactic comprising Strokes 3–5 (Fig. 4C-5) according to the width of the two branches. He hoped to determine the reasons. He hovered on the branches of these two tactics to examine the detailed views (Fig. 4C-2 and 4C-3) and found that the increment at the later tactic comprised two parts (Fig. 4C-3). One part was caused by the adjustment in the current tactic, whereas the other part was caused by adjustments in prior tactics.

The expert further clicked this tactic (Fig. 4C-8) to examine how adjustments in prior tactics caused the increments in the impact view (Fig. 5). He quickly found that the probabilities of the first stroke being with *topspin L/M* and *topspin L/B* increased (Fig. 5A). The experts commented that these two kinds of strokes' probabilities increased at the prior tactic (Fig. 4D-7). He hovered on several increased probabilities in this tactic and found in the detailed views that many of them were positively influenced by the adjustments at prior tactics (Fig. 5B). Therefore, the scoring rate after this tactic was increased by prior adjustments (Fig. 5C).

This case study demonstrates how the expert identified an unexpected return adjustment in the adjustment view (Fig. 4B-1) and browsed the adjustment's even impact dis-

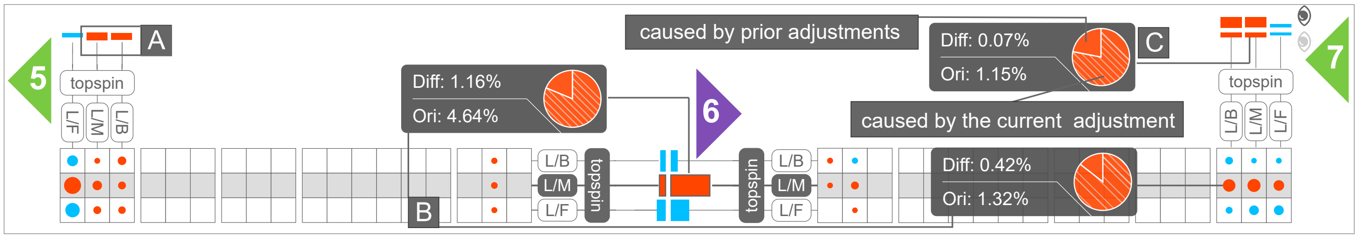


Fig. 5. Illustration of the impact view for the case study discussed in Section 5.1.2. (A) The probabilities of the first stroke being with topspin L/M and topspin L/B increased. (B) Many of the increased probabilities were partly caused by adjustments at prior tactics. (C) The increase in scoring probability after the third stroke of the tactic was partly due to prior adjustments.

tribution over tactics in the accumulation view (Fig. 4C-1). The case study also indicates how the expert interpreted the increased scoring rate at the tactic comprising Strokes 3–5 using the changed probabilities at two returns of this tactic represented by two connected matrices in the impact view (Fig. 4D).

The identified patterns and interpretations cannot be found using existing tools for table tennis data analysis, such as the software used by the experts (introduced in Section 2.2). The software cannot help users find interesting adjustments with an overview of adjustment impacts or help users interpret the impacts with intermediate details, such as how many impacts are accumulated at tactics and how the impact is generated due to the changed probabilities in two returns in a tactic. Previous visualization systems in the sports visualization area have not investigated how to visualize a sports simulation and cannot be applied directly to find out the patterns as detailed above.

#### 5.1.3 Bad Performance in Receiving Half-long Backhand after Short Forehand

This case study also analyzes matches between Ma Long and right shake-hand players. The expert selected Ma Long and right shake-hand players in the player view (Fig. 4A-3) and then browsed the adjustment view. She focused on the return adjustments that let the opponents employ strokes with *drop shot* (a stroke technique value) more frequently after Ma's serve strokes (Fig. 6A). She found that the scoring rate mostly decreased when the opponents' strokes were to short ball positions (most rectangle colors are blue, Fig. 6A-1) whereas mostly increased when the opponents' strokes were to half-long ball positions (most rectangle colors are orange, Fig. 6A-2). The expert commented that this pattern is consistent with domain knowledge. When the opponent's strokes were to half-long ball positions more frequently, Ma Long had more chances to attack and score. However, there is an outlier (Fig. 6A-3). The expert hoped to determine the reasons and clicked the entry.

The expert browsed the accumulation view and found the impacts mainly took place in the first tactic (Fig. 6B-1). She clicked the first tactic and examined how the impact was generated in the impact view. She found that the adjustment increased the probabilities that Ma Long employed *chop long* (a stroke technique value) in the third stroke (Fig. 6C-1) but the scoring probability of using *chop long* in total was decreased (the bar representing the changed scoring rate is blue, Fig. 6C-3). Meanwhile, the probabilities that Ma Long employed attacking techniques, such as *topspin*, were decreased (Fig. 6C-2). The expert commented that *short forehand* (a ball position value) is close to *half-long backhand* (a ball position value). When the opponents gave strokes to *half-long backhand* after Ma's serve strokes to *short forehand*, Ma Long was not able to react timely with attacking strokes and could only respond with *chop long* with few chances of scoring. Therefore, it decreased Ma's scoring rate to increase the opponents' strokes to *half-long backhand* after Ma's serve strokes to *short forehand*.

This case study demonstrates that the adjustment view enabled the expert to detect abnormal impact patterns (Fig. 6A-3). The accumulation view helped him find the impact at the tactic comprising Strokes 1–3 was the main cause of the abnormal impact (Fig. 6B). And the impact view helped the expert find the reasons by highlighting main changes caused by the adjustment in the second return at this tactic (Fig. 6C). This case study also demonstrates that the experts who were not involved in developing the system can use the system to find interesting impacts and interpret them.

### 5.2 Insights Obtained by the System

Using SimuExplorer, the experts verified existing hypotheses, such as "For most players, the scoring rate is decreased if we let the opponents employ strokes with *topspin* more frequently after their strokes with *topspin*," and formulated new hypotheses, such as "For particular players, such as Ma Long, it increases the scoring rate to let the opponents employ *topspin* more frequently to *long middlehand* after their strokes with *topspin*." More hypotheses and their explanations can be referred to in the supplementary materials.

### 5.3 Expert Feedback

The feedback and suggestions of the four experts are summarized in three aspects as follows.

**Usefulness.** Both the collaborating experts (experts A and B) and new experts (experts C and D) agreed that the system is useful for helping users interpret the simulated impacts of return adjustments. Experts A and B mentioned that the system can help them interpret how a return adjustment impacts the scoring rate and realize their analysis tasks. Expert A commented: "This system considers new key features of the return adjustments, i.e., combinational use of stroke technique and ball position and the tactical associations. Additionally, the system visualizes how the impacts of return adjustments are generated and

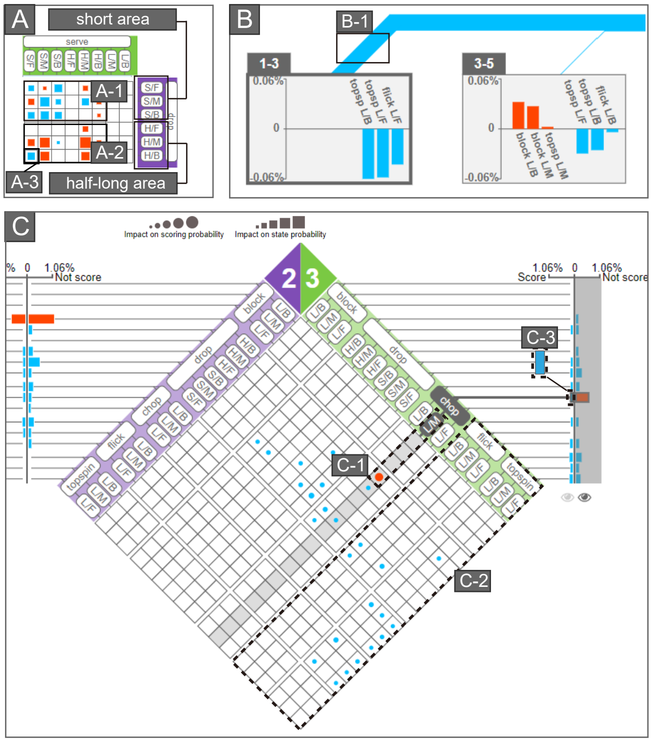


Fig. 6. Illustration of the patterns in Section 5.1.3. (A) presents the impacts of adjusting the opponents' strokes with *drop shot* (a stroke technique value) after Ma Long's serve strokes. (B) presents a selected adjustment's accumulated impacts at the first two tactics. The expert clicked the first tactic and (C) presents how the impact was generated in the second return of this tactic.

accumulated over tactics for the first time. I think it is a very advanced tool for analyzing the return adjustment impacts." Expert C mentioned that the system enables her to look at the intermediate process of how an adjustment impacts scoring rates in table tennis simulation. She had used the Markov chain model to do a simulative analysis and this system visualizes important details to help her explain the results of the model. Experts C and D both mentioned that the impact view provided them with insights into how to improve players' behaviors. In particular, experts A, B, and C thought the four views in the system are useful. Expert D thought all the four views but the player view are useful. He would like to examine the adjustment view of each player instead of browsing the player view.

**Usability.** Four experts agreed that the system is easy to learn and use on the whole. Experts A and B appreciated the impact view, which enhances understanding the transfer process of the changed probabilities. Expert B commented: "The impact view comprehensibly presents how the changed probabilities branch, merge, and finally lead to the changed scoring rate." Experts C and D could understand most designs of the system in the second phase (i.e., the introduction of the system design). Expert C thought all the views but the adjustment view are easy to learn and use. She did not learn the meaning of the adjustment view at first. Nevertheless, she thought the view is easy to use after she understood it. Expert D thought all the views but the impact view are easy to learn and use. He was not familiar with the Markov chain model and did not understand the impact view at first. However, after we explained the model and its connection to table tennis knowledge, expert D understood and successfully finished the task. He still thought the impact view is a little complex. But he thought the complexity is mainly due to the analysis task instead of the design.

**Suggestions.** Expert A suggested that the system should be enhanced to allow adjustments to multiple returns. Experts B, C, and D thought the system is comprehensive regarding the analysis tasks.

## 6 Discussion

**Lessons learned.** One lesson for pattern discovery [54] involves how to display the intermediate impacts of a player's return adjustments in an interactive table tennis rally. Initially, we planned to present the impacts at each return, and to explore their variations over returns. However, because table tennis is so highly interactive and the two players' returns appear alternately, it is hard to visually track the impacts of one player's return adjustment. To solve this problem, we grouped two consecutive returns into a tactic, a strategy based on our expert collaborators' domain knowledge. In this way, a rally can be shown as a sequence of a player's tactics, and the intermediate impacts can be visually explored.

**Limitations.** Three limitations to our study, which could be addressed in future work, are listed as follows. (1) Adjustments can be made more flexible. In our system, when we increase the probability of a return, other relevant returns are proportionately reduced to maintain a constant total probability. However, as some returns are more relevant to the increased return than others, it would be better to reduce returns differentially. (2) The system cannot support exploration of adjustments to multiple returns. Considerable efforts can be exerted to improve the system so that it can help users explore impacts caused by adjusting combinations of multiple returns. (3) The system can be enhanced to distinguish the impacts of return adjustment under different conditions, such as when there are different scores.

**Generalizability.** Game sports, such as squash, tennis, badminton, volleyball, and baseball, are most appropriately regarded as dynamic interaction processes between two parts (teams, doubles, or singles) [1]. Therefore, the Markov chain model can also simulate structures of these sports, if states are properly identified, and our system could be extended to work with these sports. The impact view can also be applied to other domains, such as those involving many-to-many relations in event sequences as discussed in Section 3.3. This view provides a layout for sequential matrices and can be extended to include more than two such matrices. Unlike the sequential matrices presented in previous studies [22], [49], our impact view proposes a new layout with sequential matrices in a straight line, a more consistent and understandable design.

## 7 Conclusion

This work investigates the problem of unfolding the simulative Markov chain model and exploring the intermediate generation and accumulation of impacts caused by return adjustments in table tennis. We worked closely with experts to determine an appropriate method of breaking a rally into tactics and identifying the intermediate impacts of return adjustments in these tactics. We further developed SimuExplorer, a visualization system that supports multi-level explorations of these impacts. We then conducted three case studies to demonstrate the usefulness and usability of our system. The main implications of this work are as follows. (1) This study investigates the problem of visually unfolding and interpreting a simulative model. (2) Our system enables experts to gain insight into the impacts of return adjustments in table tennis. Experts are able to use the system to effectively analyze matches involving top table tennis players.

## Acknowledgments

The work was supported by NSFC (62072400) , Zhejiang Provincial Natural Science Foundation (LR18F020001), and the Collaborative Innovation Center of Artificial Intelligence by MOE and Zhejiang Provincial Government (ZJU). This work was also supported by the Key Research Project of Zhejiang Lab (No.2021KE0AC02).

## References


[1] M. Lames and T. McGarry, "On the search for reliable performance indicators in game sports," *International Journal of Performance Analysis in Sport*, vol. 7, no. 1, pp. 62–79, 2007.

[2] M. I. Lanzoni, D. R. Michele, and F. Merni, "A notational analysis of shot characteristics in top-level table tennis players," *European Journal of Sport Science*, vol. 14, no. 4, pp. 309–317, 2014.

[3] T. C. Loh and O. Krasilshchikov, "Competition performance variables differences in elite and U-21 international men singles table tennis players," *Journal of Physical Education and Sport*, vol. 15, no. 4, pp. 829–833, 2015.

[4] H. Zhang, W. Liu, J. Hu, and R. Liu, "Evaluation of elite table tennis players' technique effectiveness," *Journal of Sports Sciences*, vol. 31, no. 14, pp. 1526–1534, 2013.

[5] M. Pfeiffer, H. Zhang, and A. Hohmann, "A Markov chain model of elite table tennis competition," *International Journal of Sports Science & Coaching*, vol. 5, no. 2, pp. 205–222, 2010.

[6] S. Wenninger and M. Lames, "Performance analysis in table tennis stochastic simulation by numerical derivation," *International Journal of Computer Science in Sport*, vol. 15, no. 1, pp. 22–36, 2016.

[7] 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*, vol. 26, no. 1, pp. 407–417, 2020.

[8] A. Franks, A. Miller, L. Bornn, and K. Goldsberry, "Counterpoints: Advanced defensive metrics for NBA basketball," in *MIT Sloan Sports Analytics Conference*, 2015.

[9] 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*, vol. 18, no. 11, pp. 2247–2256, 2016.

[10] C. Dietrich, D. Koop, H. T. Vo, and C. T. Silva, "Baseball4D: A tool for baseball game reconstruction &visualization," in *Proceedings of IEEE Conference on Visual Analytics Science and Technology*, 2014, pp. 23–32.

[11] H. Pileggi, C. D. Stolper, J. M. Boyle, and J. T. Stasko, "SnapShot: Visualization to propel ice hockey analytics," *IEEE Transactions on Visualization and Computer Graphics*, vol. 18, no. 12, pp. 2819–2828, 2012.

[12] M. Lage, J. P. Ono, D. Cervone, J. Chiang, C. Dietrich, and C. T. Silva, "StatCast Dashboard: Exploration of spatiotemporal baseball data," *IEEE Computer Graphics and Applications*, vol. 36, no. 5, pp. 28–37, 2016.

[13] X. Xie, J. Wang, H. Liang, D. Deng, S. Cheng, H. Zhang, W. Chen, and Y. Wu, "PassVizor: Toward better understanding of the dynamics of soccer passes," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 1322–1331, 2020.

[14] C. Perin, R. Vuillemot, and J. D. Fekete, "SoccerStories: A kick-off for visual soccer analysis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 19, no. 12, pp. 2506–2515, 2013.

[15] H. Janetzko, D. Sacha, M. Stein, T. Schreck, D. A. Keim, and O. Deussen, "Feature-driven visual analytics of soccer data," in *Proceedings of IEEE Conference on Visual Analytics Science and Technology*, 2014, pp. 13–22.

[16] D. Sacha, F. Al-amoody, M. Stein, T. Schreck, D. A. Keim, G. L. Andrienko, and H. Janetzko, "Dynamic visual abstraction of soccer movement," *Computer Graphics Forum*, vol. 36, no. 3, pp. 305–315, 2017.

[17] S. Ye, Z. Chen, X. Chu, Y. Wang, S. Fu, L. Shen, K. Zhou, and Y. Wu, "ShuttleSpace: Exploring and analyzing movement trajectory in immersive visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 860–869, 2020.

[18] X. Chu, X. Xie, S. Ye, H. Lu, H. Xiao, Z. Yuan, Z. Chen, H. Zhang, and Y. Wu, "TIVEE: Visual exploration and explanation of badminton tactics in immersive visualizations," *To appear in IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, 2022.

[19] A. G. Losada, R. Therón, and A. Benito, "BKViz: A basketball visual analysis tool," *IEEE Computer Graphics and Applications*, vol. 36, no. 6, pp. 58–68, 2016.

[20] T. Polk, J. Yang, Y. Hu, and Y. Zhao, "TenniVis: Visualization for tennis match analysis," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 2339–2348, 2014.

[21] 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*, vol. 26, no. 1, pp. 397–406, 2020.

[22] 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*, vol. 24, no. 1, pp. 709–718, 2018.

[23] Y. Wu, X. Xie, J. Wang, D. Deng, H. Liang, H. Zhang, S. Cheng, and W. Chen, "ForVizor: Visualizing spatio-temporal team formations in soccer," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 1, pp. 65–75, 2019.

[24] Y. Ishikawa and I. Fujishiro, "TideGrapher: Visual analytics of tactical situations for rugby matches," *Visual Informatics*, vol. 2, no. 1, pp. 60–70, 2018.

[25] C. Perin, R. Vuillemot, and J. D. Fekete, "A Table!: Improving temporal navigation in soccer ranking tables," in *Proceedings of ACM Conference on Human Factors in Computing Systems*, 2014, pp. 887–896.

[26] D. S. Tan, G. Smith, B. Lee, and G. G. Robertson, "AdaptiviTree: Adaptive tree visualization for tournament-style brackets," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, no. 6, pp. 1113–1120, 2007.

[27] L. Jin and D. C. Banks, "Visualizing a tennis match," in *Proceedings of IEEE Symposium on Information Visualization*, 1996, pp. 108–114.

[28] P. A. Legg, D. H. Chung, M. L. Parry, M. W. Jones, R. Long, I. W. Griffiths, and M. Chen, "MatchPad: Interactive glyph-based visualization for real-time sports performance analysis," *Computer Graphics Forum*, vol. 31, no. 3, pp. 1255–1264, 2012.

[29] J. Wang, J. Wu, A. Cao, Z. Zhou, H. Zhang, and Y. Wu, "Tac-Miner: Visual tactic mining for multiple table tennis matches," *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 6, pp. 2770–2782, 2021.

[30] J. Lan, J. Wang, X. Shu, Z. Zhou, H. Zhang, and Y. Wu, "RallyComparator: visual comparison of the multivariate and spatial stroke sequence in table tennis rally," *Journal of Visualization*, to be published, doi.org/10.1007/s12650-021-00772-0.

[31] J. Wu, D. Liu, Z. Guo, Q. Xu, and Y. Wu, "TacticFlow: Visual analytics of ever-changing tactics in racket sports," *To appear in IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, 2022.

[32] R. Cao, S. Dey, A. Cunningham, J. Walsh, R. T. Smith, J. E. Zucco, and B. H. Thomas, "Examining the use of narrative constructs in data videos," *Visual Informatics*, vol. 4, no. 1, pp. 8–22, 2020.

[33] T. Tang, J. Tang, J. Hong, L. Yu, P. Ren, and Y. Wu, "Design guidelines for augmenting short-form videos using animated data visualizations," *Journal of Visualization*, vol. 23, no. 4, pp. 707–720, 2020.

[34] M. L. Parry, P. A. Legg, D. H. S. Chung, I. W. Griffiths, and M. Chen, "Hierarchical event selection for video storyboards with a case study on snooker video visualization," *IEEE Transactions on*

*Visualization and Computer Graphics*, vol. 17, no. 12, pp. 1747–1756, 2011.
[35] Z. Chen, S. Ye, X. Chu, H. Xia, H. Zhang, H. Qu, and Y. Wu, "Augmenting sports videos with viscommentator," *To appear in IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, 2022.
[36] M. Stein, H. Janetzko, T. Breitkreutz, D. Seebacher, T. Schreck, M. Grossniklaus, I. D. Couzin, and D. A. Keim, "Director's cut: Analysis and annotation of soccer matches," *IEEE Computer Graphics and Applications*, vol. 36, no. 5, pp. 50–60, 2016.
[37] 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*, vol. 24, no. 1, pp. 13–22, 2018.
[38] M. Du and X. Yuan, "A survey of competitive sports data visualization and visual analysis," *Journal of Visualization*, vol. 24, no. 1, pp. 47–67, 2021.
[39] M. Wattenberg, "Visual exploration of multivariate graphs," in *Proceedings of Conference on Human Factors in Computing Systems*, 2006, pp. 811–819.
[40] A. J. Pretorius and J. J. Van Wijk, "Visual analysis of multivariate state transition graphs," *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, pp. 685–692, 2006.
[41] ——, "Visual inspection of multivariate graphs," *Computer Graphics Forum*, vol. 27, no. 3, pp. 967–974, 2008.
[42] F. Van Ham, H. Van de Wetering, and J. J. Van Wijk, "Interactive visualization of state transition systems," *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, no. 4, pp. 319–329, 2002.
[43] A. H. Robinson, "The 1837 maps of Henry Drury Harness," *The Geographical Journal*, vol. 121, no. 4, pp. 440–450, 1955.
[44] A. M. Voorhees, "A general theory of traffic movement," *Transportation*, vol. 40, no. 6, pp. 1105–1116, 2013.
[45] J. Wood, J. Dykes, and A. Slingsby, "Visualisation of origins, destinations and flows with OD maps," *The Cartographic Journal*, vol. 47, no. 2, pp. 117–129, 2010.
[46] Y. Yang, T. Dwyer, S. Goodwin, and K. Marriott, "Many-to-many geographically-embedded flow visualisation: An evaluation," *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 1, pp. 411–420, 2016.
[47] K. Wongsuphasawat and D. Gotz, "Outflow: Visualizing patient flow by symptoms and outcome," in *Proceedings of IEEE VisWeek Workshop on Visual Analytics in Healthcare*, 2011, pp. 25–28.
[48] R. A. Leite, A. Arleo, J. Sorger, T. Gschwandtner, and S. Miksch, "Hermes: Guidance-enriched visual analytics for economic network exploration," *Visual Informatics*, vol. 4, no. 4, pp. 11–22, 2020.
[49] J. Zhao, Z. Liu, M. Dontcheva, A. Hertzmann, and A. Wilson, "MatrixWave: Visual comparison of event sequence data," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2015, pp. 259–268.
[50] J. Lu, X. Xie, J. Lan, T.-Q. Peng, Y. Wu, and W. Chen, "BeXplorer: Visual analytics of dynamic interplay between communication and purchase behaviors in MMORPGs," *Visual Informatics*, vol. 3, no. 2, pp. 87–101, 2019.
[51] W. Kim, C. Shim, and Y. D. Chung, "SkyFlow: A visual analysis of high-dimensional skylines in time-series," *Journal of Visualization*, vol. 24, no. 5, pp. 1033–1050, 2021.
[52] Y. Wang, T.-Q. Peng, H. Lu, H. Wang, X. Xie, H. Qu, and Y. Wu, "Seek for Success: A visualization approach for understanding the dynamics of academic careers," *To appear in IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, 2022.
[53] Y. Zhang, C. Yu, R. Wang, and X. Liu, "Visual dimension analysis based on dimension subdivision," *Journal of Visualization*, vol. 24, no. 1, pp. 117–131, 2021.
[54] N. Andrienko, G. Andrienko, S. Miksch, H. Schumann, and S. Wrobel, "A theoretical model for pattern discovery in visual analytics," *Visual Informatics*, vol. 5, no. 1, pp. 23–42, 2021.

**Ji Lan** is a joint Ph.D. student of Yingcai Wu @State Key Lab of CAD & CG College of Computer Science, Zhejiang University, Hangzhou, China, and Jonathan ZHU @Department of Media and Communication, CityU University, Hong Kong. His main research interests are in visual analytics and computational social science. For more information, please visit https://lanjizju.github.io/.

**Zheng Zhou** received his PhD in Sports Pedagogy and Coaching Science from Zhejiang University, China in 2020. He works as a postdoctoral researcher in the Department of Sport Science, College of Education, Zhejiang University, Hangzhou, China. His main research interests are in analytics in sports science, and he is the level-2 table tennis player.

**Jiachen Wang** is a joint Ph.D. student of Yingcai Wu @State Key Lab of CAD & CG College of Computer Science, Zhejiang University, Hangzhou, China, and Hui Zhang @Department of Sport Science, Zhejiang University, Hangzhou, China. His research mainly focuses on visual analytics in sport science. For more information, please visit: http://www.wjc-vis.com.

**Hui Zhang** graduated from Zhejiang Normal University with bachelor's degree, Beijing Sport University with Master's and University of Potsdam in Germany for PhD. He has rich training and competition experiences and was the level-1 table tennis player, international referee, level-1 coach and the coach with A-level certificate in Germany. From 2003 to 2013, he was the professor of Shanghai University of Sport, the executive deputy dean of China Table Tennis College and director of Key Lab of Technique and Tactic Diagnosis and Analysis, State General Administration of Sports. Since May, 2014, he works in Sport Science Department, College of Education, Zhejiang University as a Professor. He is the secretary general of International Association of Computer Science in Sport (2017-2021) and International Senior Fellowship at University of Bayreuth (2018-2021) in Germany.

**Xiao Xie** is a Research Fellow at the Department of Sports Science, Zhejiang University. His research interests include data visualization, visual analytics, and human-computer interaction, with a focus on creating novel visual analysis techniques for supporting sports analysis. He received his Ph.D. degree in Computer Science from Zhejiang University.

**Yingcai Wu** is a Professor at the State Key Lab of CAD & CG, Zhejiang University. His main research interests are in information visualization and visual analytics, with focuses on sports science and urban computing. He received his Ph.D. degree in Computer Science from the Hong Kong University of Science and Technology. Prior to his current position, Dr. Wu was a postdoctoral researcher at the University of California, Davis from 2010 to 2012, and a researcher in Microsoft Research Asia from 2012 to 2015. For more information, please visit http://www.ycwu.org.