Team-Builder: Toward More Effective Lineup Selection in Soccer

Anqi Cao, Ji Lan, Xiao Xie, Hongyu Chen, Xiaolong (Luke) Zhang, Hui Zhang, and Yingcai Wu

Abstract—Lineup selection is an essential and important task in soccer matches. To win a match, coaches must consider various factors and select appropriate players for a planned formation. Computation-based tools have been proposed to help coaches on this complex task, but they are usually based on over-simplified models on player performances, do not support interactive analysis, and overlook the inputs by coaches. In this paper, we propose a method for visual analytics of soccer lineup selection by tackling two challenges: characterizing essential factors involved in generating optimal lineup, and supporting coach-driven visual analytics of lineup selection. We develop a lineup selection model that integrates such important factors, such as spatial regions of player actions and defensive interactions with opponent players. A visualization system, Team-Builder, is developed to help coaches control the process of lineup generation, explanation, and comparison through multiple coordinated views. The usefulness and effectiveness of our system are demonstrated by two case studies on a real-world soccer event dataset.

Index Terms—Sports Visualization, Lineup Selection, Design Study

1 Introduction

In soccer matches, team lineup selection is vital and a proper lineup with effective tactics can significantly increase the chance of winning [1]. Numerous examples among top leagues and national teams illustrate the importance of lineup selection. For instance, the unexpected success by the team Leicester City in the 2015/16 season of the *English Premier League* is partially attributed to excellent lineup selections that lead to its effective counter-attack tactic [2]. In recently held *EURO 2020*, a number of national teams, such as Denmark, Switzerland, and Czech, selected suitable lineups and adopted effective tactics, which helped them defeat their opponent teams that included star players [3].

Selecting a good lineup is not a trivial task. Coaches need to first choose team tactics and then choose players under various criteria and through multiple comparisons among the skills and performances of players [4]. Thus, lineup selection can be described as a problem that requires trade-offs between multiple criteria, some of which could be contradictory. Specifically, different players have various playing styles and perform differently on the aspect of offense and defense. The coach has to choose a lineup carefully by considering the synergy between team tactics and each player's offensive and defensive skills. This process usually involves extensive data analytical activities.

Recently, various computational tools have been proposed to help coaches find suitable lineups [5]. However, those approaches mainly use complicated models to generate the best lineup automatically, and often overlook some important information in lineup selection, such as team tactics. Without the knowledge of the computational processes involved in the models, coaches often find it a big challenge to understand how the results are obtained; and unable to incorporate more diverse information, these models cannot support more in-depth analysis of lineups. Therefore, such automatic lineup selection models are seldom adopted by coaches in professional soccer teams [5].

In addition, lineup selection often needs interactive analysis to try and compare different options and parameters based on real-time information (e.g., possible strategies of the opponent team). Existing methods for lineup analysis are usually weak in providing interactive tools. Although various interactive visualization systems have been available for soccer data analysis, such as positions and actions [6], formation variation [7], passing patterns [8], and migration [9], tools for interactive lineup analysis and selection are rare.

In this research, we propose a visualization system to provide a systematic solution for the lineup selection problem. The system can support the exploration and evaluation of various lineup options recommended by the model. Our efforts are made to address two main challenges. The first challenge is related to the difficulty in characterizing and modeling the complex factors that should be considered in lineup selection. In soccer matches, lineup selection is influenced by not only the performances of individual players but also the teamwork among players [10], [11], [12]. Previous studies have modeled the teamwork in basketball [13], [14] and soccer [15], [16] to help coaches select the best lineup. However, other important and heterogeneous factors of teamwork, such as the interactions with the opponent and corresponding tactics preferred by coaches, have not been considered yet [17]. It is non-trivial to integrate those complex factors into existing models. Another challenge lies in the difficulty in designing a systematic tool to assist soccer coaches to make the best lineup decision. As the experiences and preferences of soccer coaches vary significantly, the

- *A. Cao, J. Lan, H. Chen, and Y. Wu are with the State Key Lab of CAD&CG, Zhejiang University. E-mail: {caoanqi, lanjizju, chenhy7820, ycwu}@zju.edu.cn.*
- *X. Xie and H. Zhang are with the Department of Sports Science, Zhejiang University. E-mail: {xxie, zhang_hui}@zju.edu.cn.*
- *X. Zhang is with the College of Information Sciences and Technology, Pennsylvania State University. E-mail: lzhang@ist.psu.edu.*
- *X. Xie is the corresponding author.*

Manuscript received April 19, 2005; revised August 26, 2015.

factors that should be considered in the lineup selection model could be significant, so interactive control over these factors and relevant parameters is required. Coaches also need to evaluate different aspects of a selected player such as individual features and interactions with other players in matches, multiple criteria of lineups such as the total number of passing, shot, and interception, and the predicted match results of each lineup. The system should also allow visual comparison of players and lineups recommended by the model according to various criteria.

Teamed up with the domain experts, we develop a new model for lineup selection and design a system, Team-Builder, based on the model. Assisted by soccer coaches and soccer data analysts, we synthesize and operationalize various factors for lineup selection (e.g., tactical preferences and opponent information), and enhance a teamwork-based model [15] to integrate these identified factors in the model. Team-Builder, a visual analytics system, helps coaches adjust the model according to their preferences and evaluate model results. Our research makes the following contributions:

- we formalize **the characterization of the problem** of soccer lineup selection, including identification and modeling of key factors, specification of preferences, and evaluation of player and lineup candidates;
- we synthesize **some key factors** for lineup selection and developed a **teamwork-based model** to integrate these factors; and
- we develop a **visual analytics system** to support exploration and evaluation of lineup recommendations by the model.

2 Related Work

In this section, we review research on soccer data visualization and soccer lineup selection methods.

2.1 Visualization of Soccer Data

The interests of analyzing sports data with visualization techniques have increased significantly [18], [19], as seen in work on basketball [20], [21], [22], baseball [23], [24], [25], [26], racquet sports [27], [28], [29], [30], and other events [31], [32], [33]. There was considerable visualization work to analyze soccer match data from diverse aspects. For example, Perin et al. developed A table! [34] and Gap Charts [35] to visualize the temporal evolution of soccer team rankings; and Rusu et al. [36], [37] developed metaphor-based visualization to compare performances among different players with statistical indicator analysis.

Analytical methods in soccer data visualization systems have become more and more sophisticated. Involved data is no longer limited to simple statistics and includes fine-grained data of events during a match, data has been collected through videos directly, and movement trajectories of players and the ball. Soccer event data describe the spatio-temporal events that occurred in matches. Such kind of data is widely used for passing pattern analysis and player performance analysis [18]. To help the investigation of passing patterns among players, one of the most essential aspects of soccer event data analysis, researchers have developed systems like SoccerStories [38], a visualization system on player actions and spatial passing patterns, and PassVizor [8], a visual analytical system for dynamic passing patterns. Malqui et al. [39] also developed a visual analytics system to discover soccer passing strategies based on flow motifs. Compared with event data, soccer match videos and trajectory data are more detailed and include the positions of all players and the ball during the whole match. These data could support diverse analyses such as video analysis, trajectory analysis, and formation analysis [18]. To facilitate soccer match video analysis, Stein et al. [40], [41] proposed an automatic video annotation technique that can integrate visualization of player movements with videos. In addition to some methods for the computation of player movement trajectories, such as a trajectory search method proposed by Shao et al. [42], and trajectory aggregation methods by Sacha et al. [43], various visualization-based systems have been developed to support interactive analytics of player movement trajectories. Andrienko et al. used player trajectory data to show player defense [44] and team tactics [6]; Janetzko et al. [45] designed a system to analyze player performances based on extracted player trajectories; Machado et al. [46] provided a heatmap-based visualization for team formation identification; and ForVizor [7] developed a tailored flow-based visualization to represent the change of team formations and a system for a systematic analysis of the formations. Besides, visual analytics systems based on soccer trajectory data also contain other aspects such as what-if analysis for player passing decisions [47] and performance analysis based on player coordination [48].

Although these studies can help soccer data analysis one way or another, few could support visual analytics for soccer lineup selection, because of the difference in analysis tasks.

2.2 Soccer Lineup Selection

Soccer lineup analysis has attracted extensive research attention, especially on the aspects of formation detection [49], [50] and evaluation [51]. In particular, lineup selection is vital in soccer matches, and an increasing number of studies in this field have emerged during recent years [5]. Some research has provided models to find the best soccer lineup by aggregating the performances of individual players. Boon and Sierksma [1] pioneered a study to solve the problem of selecting an optimal soccer lineup. They estimated the weights of individual player performances on different positions, and selected the optimal lineup by maximizing the total performance value. Tavana et al. [52] further defined soccer lineup selection as a multi-criteria decision-making (MCDM) problem, and let the important weights of player performance attributes be controlled by users, such as coaches who have better knowledge about players and positions. Based on the definition of the lineup selection problem, they proposed a fuzzy inference system and a linear programming model to maximize the overall lineup score. Similarly, Ozceylan [4] also modeled soccer lineup selection as an MCDM problem, and maximized the sum of player value based on player property weights under multiple performance criteria. In addition to MCDM-based methods, other approaches were also proposed for selecting soccer lineups according to player individual performances. For instance, Merigó and Gil-Lafuente [53] presented a soccer player selection model based on business decision-making methods, and Al-Shboul et al. [54] utilized neural networks to decide the best lineup with the highest winning rate.

Soccer is a team sport, so teamwork is an essential aspect when coaches select players [10], [11], [12]. However, this team factor was largely overlooked by the methods discussed above. To integrate this important factor in lineup selection, Beal et al. [15] constructed a framework to quantify direct interactions among players from the same team with network-based metrics and found the best team using the mixed-integer programming method. Bransen et al. [16] improved this framework with a data-driven approach to calculate interactions by the change of goal probabilities. However, these models do not consider factors related to the opponent and the corresponding tactical inputs from coaches. The lack of such factors limits their applications, because lineup selection can vary from opponent to opponent and different coaches may have different opinions when competing with the same team [52].

To address the limitations and challenges in the visual analytics of soccer lineup selection, we propose a teamwork-based method that considers the performances of individual players, information on the opponent team, and the inputs from the coach.

3 Background and System Overview

In this section, we first introduce the relevant terminology in soccer analysis and the data used in this research. Then, we describe our work to develop the design requirements assisted by domain experts.

3.1 Background

Soccer is a highly dynamic team sport. In a match, two teams attempt to shoot the soccer ball into the goal of the opponent, and the team with more goals wins the match. Each team has eleven players on the field, and they assume one of the four roles: goalkeeper, defender, midfielder, and forward.

An action is an event started by a player who has the ball under control or tries to take over the control of the ball. It is a fundamental element in soccer event analysis [55]. The details of the event attributes are shown in Table 1 [55], [56]. The actions could be further divided into offensive and defensive actions according to the event type. We refer the actions with event types of foul, tackle, interception to defensive actions and those with other types to offensive actions. Generally, the offensive actions indicate how the players take or pass the ball to create a goal, and the defensive actions aim to take back the possession of the ball from the opponent team.

A phase is a series of consecutive actions performed by players in the same team with the ball under control [57]. A phase starts when a player takes control of the ball, and ends when the ball is lost to the opponent. During a phase, the team controlling the ball tries to win a goal by successive actions, such as passing, dribbling, and shooting.

A tactic refers to a set of similar actions or phases that leads to the same tactical effects. We use the definition of an offensive tactic as a cluster of phases in similar spatial regions with a purpose to drive the ball to the goal of the opponent [57]. In soccer event analysis, defense could be reflected from defensive actions to intercept the passing of the opponents or take the ball directly. Therefore, we use the definition of a defensive tactic as a set of actions in similar spatial regions to defend the goal [17].

An offensive interaction based on soccer events data is composed of two consecutive actions by two players in the same phase [15]. It is the basic unit to evaluate how players cooperate through the movement of the soccer ball. For instance, if a player is surrounded by several opponents and has difficulty in taking further actions, the player will pass the ball to a teammate in a better position. If an offensive interaction is successful, the chance to score a goal can be increased. Therefore, offensive interactions between players are an essential indicator in soccer lineup selection [15], [16].

A defensive interaction could be defined as two consecutive actions of players from different teams. Specifically, a defensive interaction is composed of the last action of the previous phase and the first action of the next phase. In soccer event data, the confrontation between two teams can be reflected by the soccer ball transiting through defensive actions [17]. Similar to offensive interaction, a successful defensive interaction will increase the goal chance of a team, and also decrease the goal chance of the opponent.

A lineup is the eleven players of a team selected by the coach to appear in a match. Before a match, the coach and staff members will analyze the offense and defense of the opponent team, and select well-performed players with the consideration of such opponent information. Considering the importance of teamwork, the lineup also considers how well those selected players can work together.

3.2 Data Description

The data used in our analysis is from the soccer event dataset published by Pappalardo et al. [56]. The dataset contains not only spatio-temporal event attributes, but also contextual data, such as players involved in each event, player team, and match information. The primary event attributes involved in the analysis are shown in Table 1.

The data we used for analysis contains all matches in the five major European soccer leagues in the season 2017/18. In detail, the data consists of 380 matches from Spanish *La Liga*, 380 matches from *English Premier League*, 380 matches from Italian *Serie A*, 306 matches from German *Bundesliga*, and 380 matches from French *Ligue 1*. The total number of matches is 1,224, including 98 teams and 4,229 players.

3.3 Requirement Analysis

We worked with a team of three domain experts to learn about the methods for lineup selection. The expert team included a coach of a professional soccer team, a professor of physical education who is a senior sports analyst with decades of experience, and a doctoral student in physical education who was a professional soccer player in a top national-league team.

Our goals are multi-fold. First, we wanted to characterize the problem of soccer lineup selection through meetings and interviews with the experts. Second, we expected to develop a lineup selection model based on what we have learned from the experts. Finally, we sought help from them

TABLE 1
The event attributes

Attribute	Description
Event Type	A technique that a player takes to deal with the ball (*pass, cross, free-kick, dribble, shot, foul, tackle, interception, offside, etc.*).
Event Result	Result of the event (*succeeded or failed*).
Time	Time as the event occurred in a match.
Positions	The starting and ending coordinates of the event in the field.
Player	The player who started the event.
Player Role	The role of the player (*goalkeeper, defender, midfielder, or forward*).
Team	The team of the player.
Match ID	The index of the match.

in developing design ideas and design requirements for our visual analytics system.

Characterization of the problem. We reviewed relevant literature [1], [4], [15], [16], [52], [53], [54] to summarize the current problem characterization of soccer lineup selection and held a series of meetings with the experts. During the meetings, we mainly discussed two questions: whether the problem characterization in literature is consistent with their experiences, and how lineup selection is made in real soccer matches. We also had interviews with each expert to learn about their own preference and criteria used in lineup selection. We recorded the discussion process in the meetings and the answers of each expert in the interviews. Based on data collected from the meeting discussions and interviews, we drafted the initial requirements for system design.

Development of the lineup selection model. We applied a series of models and discussed if the model results meet their expectations and how to improve the models with the experts. They compared these models, and stated that teamwork-based models are more practical and straightforward than deep learning models, because the results of the former match their domain experience better. Furthermore, the experts suggested that the current teamwork-based models could be improved by considering other essential information such as the opponent and tactical choices of coaches. Therefore, we constructed a teamwork-based model by integrating opponent and tactical information.

Iteration of visual design. After the model development, we designed a visual design prototype based on the initial requirements and our model. The prototype included tactic detection and automatic lineup recommendation according to preferred tactics of the users. We presented the prototype to the experts for feedback. They commented that explaining why a player is selected by the model is also demanded by coaches. Besides, comparison among multiple lineup selection results and what-if analysis would help them decide the most suitable lineup. Their suggestions helped us improve design requirements and the prototype. After several rounds of iterations, design requirements and the prototype were stabilized. The timeline of the design iteration process is provided in our supplementary material.

Design requirement development. Our data indicate that a system to support lineup selection need to provide tools for tasks at three levels: *tactic-level exploration*, *player-level investigation*, and *lineup-level comparison*.
Tools for *tactic-level exploration* aim to support coaches to identify tactical preferences when facing a given opponent by integrating their domain knowledge of lineup selection.

T1 *Supporting the exploration of different categories of tactics*. When selecting a lineup before a match, coaches need to obtain an overview of tactic categories of their team and those of the opponent. Providing necessary tactic information can help coaches comprehend the characteristics of both teams and further decide the tactics for a match.

T2 *Supporting the exploration of available tactics and their effectiveness according to spatial regions when facing a given opponent*. Coaches tend to analyze the tactics of the opponent and prepare corresponding tactics on the aspects of offensive and defensive. For instance, if the opponent team tends to attack from the left side, coaches will choose players who are good at defending on the right. This information can provide references for the selection of players who are suitable for attacking the weakness and defending the strike of the opponent team.

Tools for *player-level investigation* can help understand the player selection process and interpret the selection results.

P1 *Supporting the comparison of player performances under multiple criteria*. After deciding the tactics for the match, coaches would like to select several players who performed well within the selected tactics as the core players in the lineup. They need to compare the performances of these players under different criteria. Thus, it is necessary to support efficient comparison among players to help coaches select suitable core players.

P2 *Supporting the explanation of player selection in the lineup result*. The lineup selection model can recommend players based on the coaches' preferences. Coaches require to know why a player is selected by the model to determine whether it is reasonable to include the player in real matches. The lineup selection result is associated with players' individual performances and interactions with other players. The tool needs to provide these details.

Tools for *lineup-level comparison* assist to obtain information of multiple results and to decide the most suitable lineup.

L1 *Supporting the comparison of expected match results and reliability of different lineups*. Coaches may generate several candidate lineups under different preferences and select one of them considering both expected match result and reliability. The expected match result is the probability of winning/drawing/losing the match, and the reliability means the appearance time of players in previous matches. Tools should be provided to help compare these metrics and decide on the most appropriate lineup.

L2 *Supporting the comparison of predicted statistical indicators of different lineups*. When the lineup used for a match is decided, coaches demand to obtain the predicted statistical indicators of the lineup and compare them with those of other candidate lineups. These indicators could help coaches evaluate whether the tactical characteristics of the lineup meet their expectations. Such detailed knowledge could also facilitate the choice of suitable match strategies such as build-up and counter-attack.

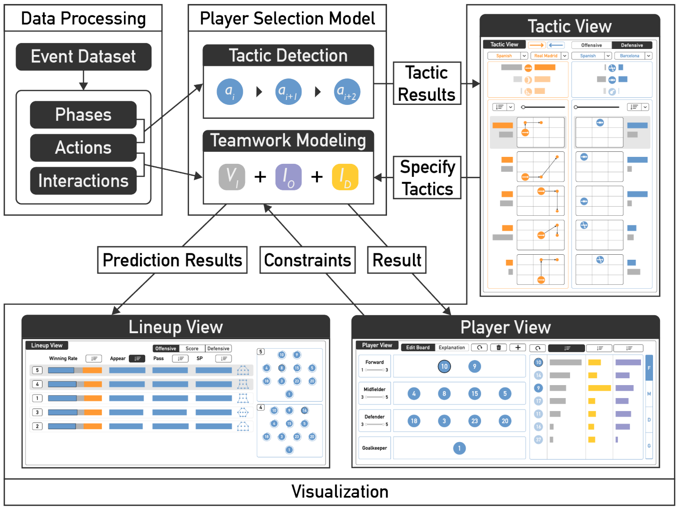


Fig. 1. Major components of the system and their relationships. The system has three components: the data processing component, the player selection model component, and the visualization component.

3.4 System Overview

We develop Team-Builder, a web-based application to support visual analytics of soccer lineup selection. The system has three components: a data processing component, a player selection model component, and a visualization component (Fig. 1). The data processing component supports the extraction of phases from the raw event dataset and the calculation of the value for each action and interaction. The player selection model component establishes the teamwork-based model for the selected target team with the data from the data processing component. The visualization component includes the user interface of the system for interactive lineup selection and consists of three views: the tactic view, the player view, and the lineup view. The system was implemented with MongoDB for the data processing component, Flask in Python for the player selection model component, and React for the visualization component.

4 Teamwork-Based Model for Player Selection

In this section, we describe our teamwork-based model to generate lineups for soccer matches.

4.1 Task Definition

The major task of soccer lineup selection is to identify eleven players for a certain match from all players in the team. We define the set of candidate players as $P = \{p_1, \ldots, p_{n_P}\}$ (n_* indicates the amount of the elements in set $*$) and a lineup as one of the subsets of P with eleven elements.

Soccer lineup selection is a complex process. We simplify the process and only consider the most essential factors in lineup selection. According to previous work [1], we define the most suitable lineup as the one with the best team performance. Based on the definition, the lineup selection process can be modeled as an optimization problem of choosing eleven players with the maximum value of a specific team performance metric [1]. In soccer matches, the team performance can be estimated by the historical match data and the physical and mental conditions of players. We further simplify the process by estimating the team performance only using the historical match data.

Several metrics have been proposed to select the optimal lineup, including the match winning rate [54], the total performance score [1], [4], [52], and the teamwork value [15], [16]. During the cooperation with our experts, we found that the teamwork value is more practical than other metrics because interaction among the players is also important for a match. Therefore, we use teamwork value as the team performance metric that is maximized in our model.

The teamwork value among players refers to the composition of the total individual performance value and the total interaction value of all players in the lineup. In soccer matches, the interactions among the players could be divided into direct interactions completed through ball transition between two players, and indirect interactions accomplished by the coordinated movements of the players [15]. We only consider the direct interactions between two players because they are the most basic and important parts of player teamwork [15]. Thus, in the calculation of the teamwork value, a soccer match can be treated as a sequence of player actions $A = \{a_1^p, \ldots, a_{n_A}^p\}$, where $p \in P_M$ is the player who performs the actions and P_M is the set of all players involved in the match (blue & orange objects in Fig. 2).

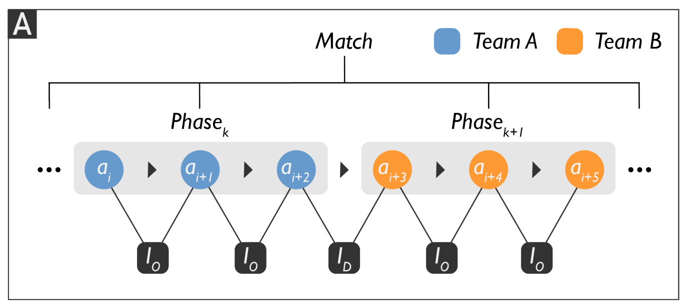


Fig. 2. The structure of a soccer match. a_i refers to an action, I_O an offensive interaction, and I_D a defensive interaction. The gray objects illustrate that in a soccer match, a new phase would start when the ball possession has changed.

Some teamwork-based models have been proposed for soccer lineup selection [15], [16]. However, those models have some critical limitations in real-world practices.

Neglecting spatio-temporal information. The original measurements of player interaction $I(p_i, p_j)$ considered all the consecutive two actions $(a_k^{p_i}, a_{k+1}^{p_j})$ of a certain pair of players (p_i, p_j) in the same team and aggregated the values directly (black objects in Fig. 2). However, these measurements do not reflect the tactical preference of coaches (e.g., wing-attack, mid-attack) because the spatio-temporal information of phases is lost. Valuing teamwork without spatio-temporal information is inadequate.

Ignoring the opponent of the team. Player interactions with the opponent indicate defensive features between players from different teams, and are also an important factor when coaches select the lineup. However, previous models did not distinguish teamwork values with opponents and provided the same solution for diverse opponents. Therefore, it is insufficient to generate solutions in practical matches.

To address these two limitations, we develop a new lineup selection model that considers the spatio-temporal information and the opponent. As mentioned before, our model only considers part of the problem and is just a step to understand the problem better. For the first limitation, we integrate a tactic detecting model into the calculation of teamwork value to identify important interactions. The tactic detecting model could find different tactics $T = \{t_1, \ldots, t_{n_T}\}$ by spatio-temporal similarities. Coaches can choose preferred tactics and the weights of corresponding interactions will be increased in the calculation of the teamwork value. With such a method, players who perform well in the selected tactics will be more likely to be chosen. To deal with the second limitation, we divide the interactions among players into offensive interactions I_O and defensive interactions I_D (black objects in Fig. 2). In this way, the teamwork value is associated with the actions of the opponent.

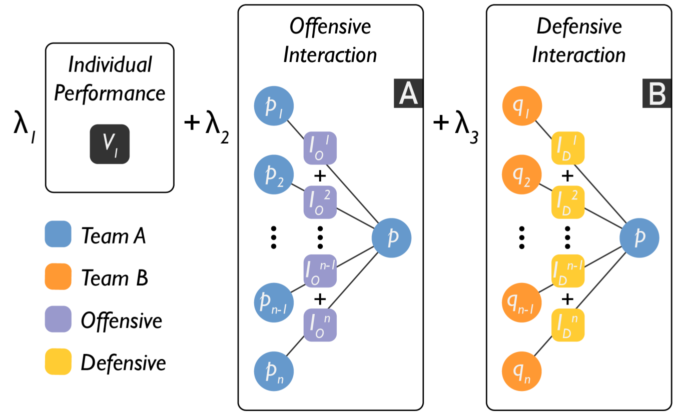


Fig. 3. The calculation of the teamwork value of Player p. p_i refers to a player in the same team and q_i a player in the opponent team. I_O^k and I_D^k represent offensive and defensive interaction value between two players.

4.2 Tactic Detection

When selecting a soccer lineup, coaches usually pay attention to the tactics used by the target team and its opponent to evaluate the strengths and weaknesses of the two teams. Such information could provide vital references for coaches to select appropriate players to confront the opponent. Thus, we utilize tactic detection models to extract offensive tactics [57] and defensive tactics [17] used by a specific team. The detailed detection process of offensive and defensive tactics is provided in our supplementary material. In our lineup selection model, coaches could select the offensive and defensive tactics they desire to use, and the value of interactions included in the chosen tactics would be allocated with an elevated weight.

4.3 Teamwork Value Evaluation

Both the individual player performance and the interactions among players are essential to lineup selection. Compared with the interactions among players in the same team, the interactions between two teams indicate how the control of the ball transits, and can reflect defensive tactics. Thus, we decide to take these features to measure the teamwork value of a lineup. Based on the previous work [15], [16], we construct a model for teamwork value calculation as follows:

$$V = \lambda_1 \cdot V_I + \lambda_2 \cdot I_O + \lambda_3 \cdot I_D, \tag{1}$$

where $\lambda_1 + \lambda_2 + \lambda_3 = 1$. V_I denotes the total individual performance value, I_O and I_D refer to the total offensive interaction value and the total defensive interaction value, respectively. λ_1, λ_2, and λ_3 are the weights of different features, which are estimated from historical match data.

As mentioned before, we only take the direct interactions into the calculation of the teamwork value. The teamwork value of a certain player p is shown in Fig. 3, composed of the individual performance value, the offensive interaction values with the teammates selected in the lineup (Fig. 3A), and the defensive interaction values with the opponents (Fig. 3B). Here we introduce the definition of those features. For simplicity, we use $P = \{p_1, \ldots, p_{n_P}\}$ to represent the set of all candidate players in a team.

Individual performance evaluation. For a player $p_i \in P$, we use the individual performance value $V_I(p_i)$ to indicate how well the player performed from the aspect of personal ability. According to existing research, individual performance can be measured by abstract statistic indicators such as passing network attributes [15] and the contribution to goal probability [55], [58]. In our case, we choose the approach that is based on the changes of goal probabilities for the player's own and opponent teams [55]. We calculate the performance value of each action on reward and risk. The calculation process is provided in our supplementary material. The individual performance value of a player is calculated by aggregating the performance value of all actions of the player.

Offensive interaction value. For a pair of players (p_i, p_j), where $p_i, p_j \in P$, we use the offensive interaction value $I_O(p_i, p_j)$ to indicate the effectiveness of player interactions in offense (Fig. 3A). Generally, coaches treat each phase as a complete offense process and the basic unit for team offense analysis [57]. Therefore, the interactions among players in the same team during the same phase could be regarded as offensive interactions (black objects in Fig. 2). On the basis of the previous works, $I_O(p_i, p_j)$ is the aggregation of all interactions between p_i and p_j. The equation is as follows:

$$I_O(p_i, p_j) = \sum_{k=1}^{n} (w_k \cdot I_O^k(p_i, p_j)) \cdot \frac{90}{M(p_i, p_j)}, \tag{2}$$

where k refers to the index of the current interaction in the set of all interactions between p_i and p_j, n represents the total number of interactions between p_i and p_j, $I_O^k(p_i, p_j)$ means the offensive interaction value on the index k, $M(p_i, p_j)$ is the number of minutes that p_i and p_j have played together, and w_k is the weight of the corresponding interaction. $I_O^k(p_i, p_j)$ is gained by the addition of the values of the two consecutive actions involved in the interaction. The calculation of the action value is the same as that for the individual performance value. If a tactic is selected, all interactions that belong to the tactic will have an elevated weight $w_k > 1$. The value is normalized by 90 minutes, the regular time of a soccer match, to avoid ignoring players who seldom played together but interacted effectively.

Defensive interaction value. For a pair of players from different teams (p_i, q_j), where $p_i \in P$, $q_j \in Q$ (Q is the set of players of another team), we use the defensive interaction value $I_D(p_i, q_j)$ to evaluate player defense (Fig. 3B). The beginning of each phase implies a successful defense. Thus, a defensive interaction can be described by the last action in the former phase and the first action in the latter phase (black objects in Fig. 2). To gain a comprehensive defensive interaction feature, we calculate the defensive interaction value with all historical match data of the target team. To cope with the specific opponent team in the match, we assign an elevated weight to the interactions with the same opponent in previous matches to indicate the importance of those interactions. Similar to the offensive interaction value, the equation of defensive interaction is defined as follows:

$$I_D(p_i, q_j) = \sum_{k=1}^{n}(w_k \cdot I_D^k(p_i, q_j)) \cdot \frac{90}{M(p_i, q_j)}, \qquad (3)$$

where $I_D^k(p_i, q_j)$ indicates the defensive interaction value between p_i and q_j on the index k. The calculation of defensive interaction value is the same as offensive interaction value.

4.4 Solving the Optimization Equation

We solve the optimal lineup by maximizing the teamwork value defined in Equation 1 through integer linear programming. Specifically, we construct the optimization equation based on the work of Beal et al. [15] and Bransen et al. [16]. We simplify the optimization process by only considering the most important constraints, including the total player number, the team formation, and the players manually included by coaches and who cannot appear in the match. Please refer to our supplementary material for details.

4.5 Model Evaluation

Our model was evaluated with the soccer event dataset used in our analysis, including all matches in the 2017/18 season from the five top European leagues [56]. In the experiments, we split the matches in the dataset into two half-seasons for training and testing. Besides, we evaluated the performance of our model on the matches from the different five leagues.

In the first experiment, we evaluated our model with the same players in lineups in real matches, and compared the results with the previous model that does not consider the defensive interaction value [16]. We measured both models by comparing how many players in each lineup generated by a model match the players in real lineup. Specifically, we trained the two models with the match data in the first half-season, and generated lineups with the two models for each match in the second half-season. We counted the number of players who are both in the generated lineup and the real lineup to evaluate the models. Fig. 4 shows the comparison of two models in terms of the averages of the same players in model-generated lineups and real. As the result shows that the lineups generated by our model (green objects in Fig. 4) are closer to real lineups chosen by coaches than those by the previous model (red objects in Fig. 4). This implies that when selecting the lineup, coaches consider not only the individual performance and interactions with teammates, but also interactions with opponent players. Thus, considering defensive interactions in the calculation of the teamwork value can help coaches select better lineups in real matches.

In the second experiment, we evaluated that teamwork value is an effective metric to measure team performance. Please refer to our supplementary material for details.

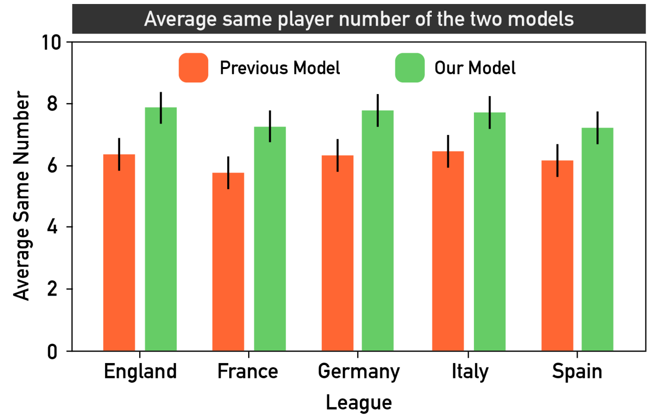


Fig. 4. The evaluation results in the first experiment. The red parts represent the previous model and the green parts represent our model.

5 Visual Design

In this section, we describe the design of visualization and interaction tools in our system based on the previous requirement analysis.

5.1 Overview of Visual Design and User Interface

According to the requirements, we design a tactic view for tactic exploration (**T1**, **T2**), a player view for lineup generation and player investigation (**P1**, **P2**), and a lineup view for solution comparison (**L1**, **L2**).

The workflow of the system is as follows. In the tactic view, users could select the target team and its opponent, and obtain an overview of all tactics used by both teams to specify essential tactic categories (**T1**) (Fig. 5A1, A2). Then, users could explore confrontation tactics in the selected categories through ranking and filtering tools to choose the desired tactic combination (**T2**) (Fig. 5A3, A4). The player view provides a candidate player list to enable users to compare different players on the individual performance and interactions with other players (**P1**) (Fig. 5B2). Based on the comparison, users can include or exclude certain players in the lineup. After the tactics and players are set, users could acquire the generated lineup and further adjust it in the lineup edit board of the player view (Fig. 5B1). The player view also contains an explanation component to explain why the players are selected in the generated lineup (**P2**) (Fig. 7). To identify the best lineup under multiple criteria, users can record lineups and compare them in the lineup view (**L1**, **L2**) (Fig. 5C). To assist the analysis, we also embed customized diagrams (Fig. 6) in the system to encode team tactics.

We use blue and orange color to code the target team and its opponent team, and purple and yellow color to code the offensive and defensive interaction values through the whole user interface.

5.2 Tactic View

The tactic view is composed of a confrontation tactic list to present all tactics used by the target team (Fig. 5A2, A4) and its opponent (Fig. 5A1, A3). The tactics are illustrated by tactic diagrams with tactic category and spatio-temporal context (Fig. 6). In this view, users can directly select the

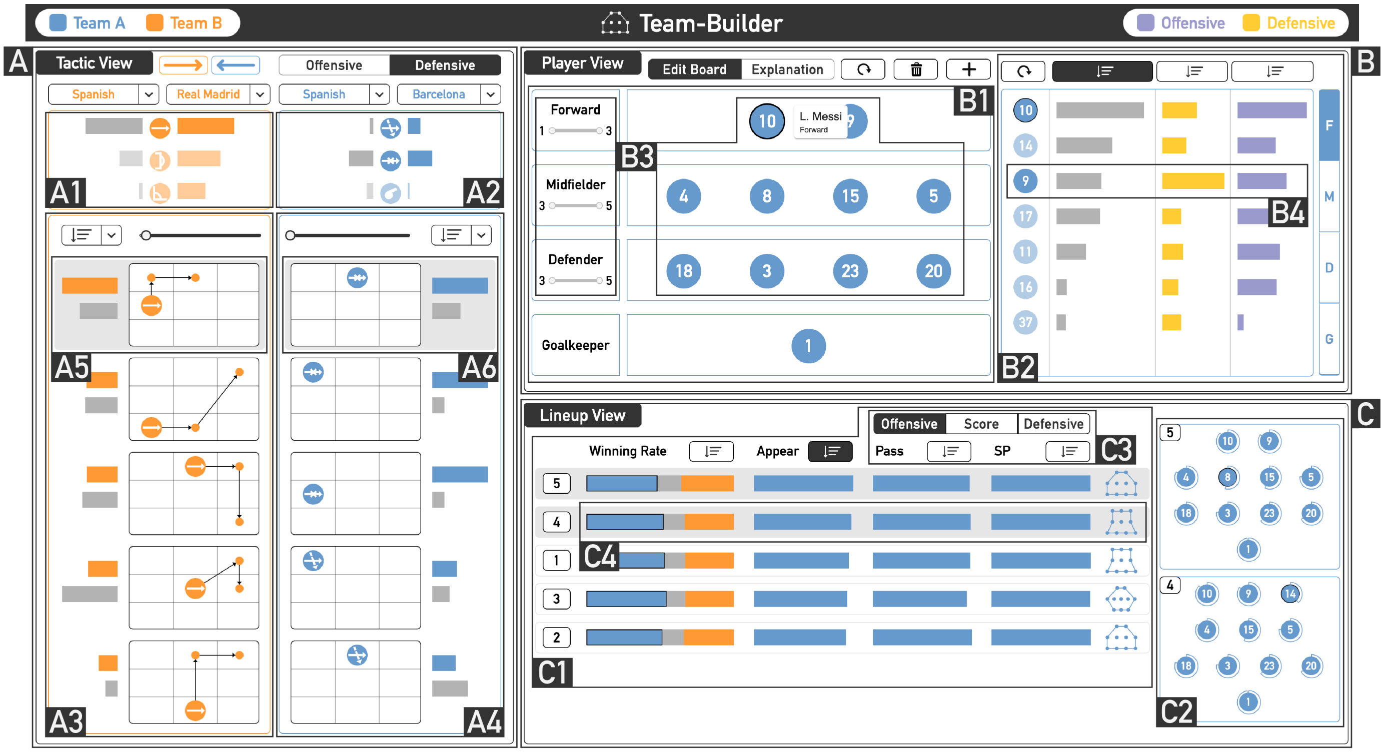


Fig. 5. System user interface. The interface consists of three views: a tactic view (A), a player view (B), and a lineup view (C). The tactic view provides confrontation tactic lists (A1, A2, A3, A4) to navigate tactics used in the lineup. The player view contains a lineup edit board (B1) for lineup generation, a candidate player list (B2) for player constraint identification, and an explanation component for comprehension of the reason of the selection of a player. The lineup view includes a candidate lineup list (C1) and lineup thumbnails (C2) for comparing multiple lineups.

target team and its opponent with the drop lists, filter tactics with essential tactic categories (**T1**), and navigate the tactics to be applied for the match through ranking by their usage rate and success rate (**T2**).

Tactic diagrams. To facilitate exploration and comparison among tactics (**T2**), we design tactic diagrams for offensive and defensive tactics in soccer matches, respectively (Fig. 6).

- ***Pitch division.*** When describing soccer tactics, the pitch could be divided into nine spatial regions, including the defensive third, the middle third, and the attacking third [17]. Therefore, we use this division method to simplify the tactic representation in our design (Fig. 6B).
- ***Offensive tactic.*** An offensive tactic is a cluster of phases on similar spatial regions. To visualize an offensive tactic, we first aggregate all phases in the tactic to an abstracted sequence that can represent the spatio-temporal feature of the tactic [43] (Fig. 6A). Then, based on the pitch division, we use points to represent the actions and arrows to illustrate the order of actions. The points are placed in the regions where the actions occurred. Therefore, an offensive tactic is visualized as a sequence of actions on the pitch (left part in Fig. 6B).
- ***Defensive tactic.*** A defensive tactic is a group of defensive actions that occurred in the same spatial region. We use a point in the particular region to represent the spatial region of a defensive tactic (right part in Fig. 6B).
- ***Icon.*** Our experts state that the offensive tactics could be classified as corner, simple free-kick, and simple pass, and the defensive tactics as tackle, interception, and foul. It is difficult to encode multiple tactic categories with color or other visual channels. Therefore, we use icons to represent different tactic categories [8] (Fig. 6C, D). The tactic category is decided by the type of the tactic's first action. Thus, we put the icon on the first action of a certain tactic to indicate the category it belongs to.

Justification. Previous work has provided representations for spatio-temporal events, such as sequence-based visualization and heatmap, but they cannot help users effectively perceive the relative positions and the order of actions in the same view [59], [60], [61]. Thus, we decide to place the actions on their occurred positions in the pitch and link them with arrows to show the action order. Moreover, coaches concern about how to specify tactics to attack the defensive weakness and defend against the attacking of the opponent. Such confrontation involves offensive and defensive tactics from different teams in the same spatial region. Thus, we apply the same pitch division to offensive and defensive tactics to facilitate the matching of confrontation tactics.

Confrontation tactic list. The confrontation tactic list includes two sub-lists that correspond with the offensive and defensive tactics of both teams (Fig. 5A1, A2, A3, A4). When users focus on the offensive tactics of the target team, the defensive tactics of the opponent team are shown simultaneously, or vice versa. We place two sub-lists side by side to represent the confrontation between two teams.

The usage and success rates are widely used in tactic style analysis. The usage rate is calculated by dividing the number of all phases by the total number of phases in the tactic, and the success rate is defined as dividing the number of phases in the tactic by the number of phases that created a shot. The usage and success rates could indicate the tactical style and the strength and weakness of a team, respectively.

The distributions of the usage rate and success rate among tactic types of the team and its opponent are most important to the understanding of the tactical styles of both teams (**T1**). Therefore, each sub-list contains distribution bar charts to present the usage rate (the gray bars) and success rate (the colored bars) of all tactic categories as an overview of tactic style (Fig. 5A1, A2). The detailed tactics are listed under the distribution bar charts. Coaches usually pay attention to the usage rates and success rates to choose confrontation tactics when facing a given opponent (**T2**). Thus, each row of the list contains the tactic diagram, the usage rate (the gray bars), and the success rate (the colored bars) (Fig. 5A5, A6). The usage rate and the success rate in the list are also encoded by bar charts for efficient ranking and filtering.

Interaction tools. Users can interact with the tactic view in various ways. The follow are some commonly used tools provided by the system.

- *Switching tactics.* Users can switch between offensive and defensive tactics of the target team with the switch button on the top right corner of the view. Meanwhile, the tactics of the opponent team will also be switched.
- *Selecting with categories.* Users can click the icons in the distribution bar charts to select categories of tactics that appearing in the list (Fig. 5A1, A2).
- *Sorting and filtering.* Users can sort the tactics by one of the usage and success rates, and filter by the other rate.
- *Selecting tactics.* Users can click the tactics of the opponent team in the left list to filter the confrontation tactics of the target team in the same spatial region. Then, the tactics of the target team in the right list can be selected for different player roles (Fig. 5A5, A6).
- *Unfolding details.* Users can hover on the icons and the bar charts to show the tactic category that the icon represents and the value of usage rate or success rate.

5.3 Player View

After selecting tactics that are expected to be used, users can generate the optimal lineup under current tactical preference in the player view. The player view consists of a lineup edit board (Fig. 5B1), a candidate player list (Fig. 5B2), and an explanation component (Fig. 7). In this view, users can explore player performances and select players included or excluded in the lineup in the candidate player list (**P1**). Then, users can obtain the generated lineup and further adjust it in the lineup edit board. In the explanation component, users can explore the interactions among the players to explain why the players are selected by the generated lineup (**P2**).

Candidate player list. The candidate player list provides a sortable list of player metrics for users to specify core players who need to be included and players who cannot appear in the match (**P1**) (Fig. 5B2). We present the players with the same role in the list because coaches need to compare them when selecting core players. Each item in the sortable list consists of an icon that represents the player, and player metrics including the individual performance value, the total offensive interaction value, and the total defensive interaction value (Fig. 5B4). All the metrics are symbolized as bar charts for simplicity and effective comparison. The maximum length of the bar charts encodes the learned weight of the metric. The icons of the players selected by users are highlighted with black borders, and those of the excluded players are indicated with the less-saturated color. Players in the list can be sorted by each kind of metric because coaches expect to select players according to different criteria.

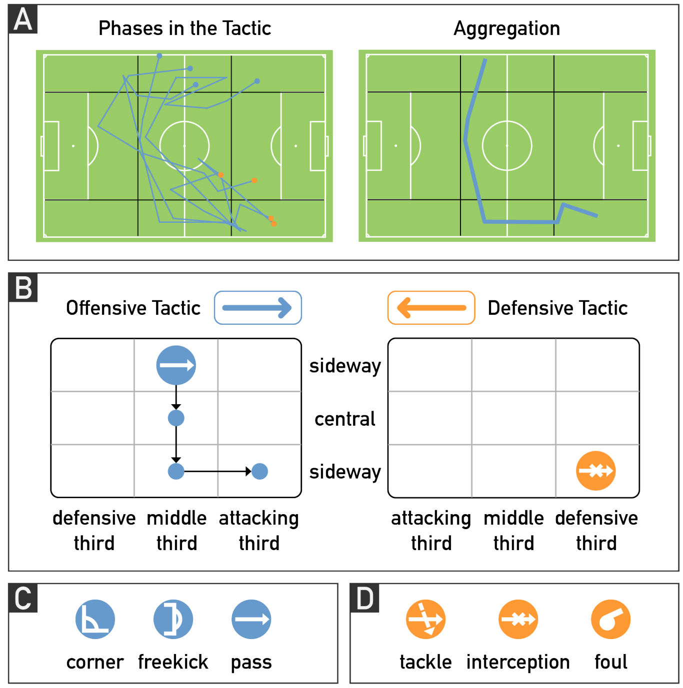


Fig. 6. Visualization of tactics. (A) illustrates the simplification of offensive tactics. (B) shows the design of the tactic diagram. (C) and (D) are icons for offensive and defensive tactic categories [8].

Lineup edit board. The lineup edit board shows the lineup result solved by the optimization equation or adjusted by users (Fig. 5B1). As the formation is an essential constraint when coaches select players, we divide the players in the lineup by their roles in formation, i.e., goalkeeper, defender, midfielder, and forward, and place the players with the same role together (Fig. 5B3). The lineup result is illustrated by actual positions in the formation for easy comprehension.

Explanation component. The explanation component shows the detailed interaction values of each player (**P2**) (Fig. 7). The interaction values among the players are represented by a matrix and encoded by the area of rectangles. The players of the target team are illustrated by the rows of the matrix and are ranked by the total value of the row (Fig. 7B2). The matrix contains two modes to show offensive and defensive interaction values (Fig. 7A1). The columns of the matrix indicate the teammates in the offensive interaction mode, or the opponents in the other mode. In the offensive interaction mode, the rows and columns present the passers and the receivers during the interactions. The players on the columns are initially summarized by player roles and can be unfolded to show the players in a certain player role (Fig. 7B1, C1). When the players on the columns are unfolded, the individual performance and total defensive interaction values of each player are also illustrated by bars (Fig. 7C3). Those players selected in the lineup are highlighted with the more saturated color for their icons (Fig. 7B2, C1).

Justification. We present the relationships among the players with a matrix, rather than a node-link diagram, to

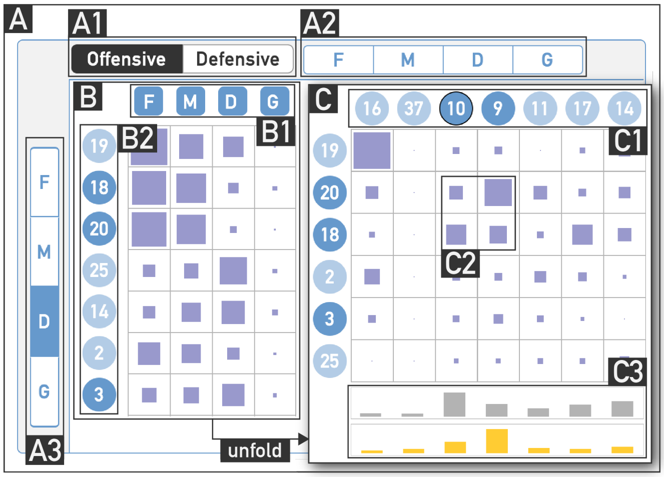


Fig. 7. The explanation component. (A) includes the switch buttons to switch among different modes. (B) illustrates the folded offensive interactions under the receiver mode. (C) shows the offensive interactions unfolded by forwards under the receiver mode.

avoid visual clutter. Coaches require to compare the selected players with the unselected players in the same role to explain why the players are selected by the model. Thus, rather than presenting all candidate players, the row and the column of the matrix only illustrate players in the same role.

Interaction tools. Users can interact with the player view in various ways. The follow are some commonly used tools provided by the system.

- *Specifying players.* Users can click the player icons in the candidate player list to specify the players who need to be included and those who cannot appear in the lineup.
- *Adjusting constraints.* Users can drag the sliders in the lineup edit board to adjust the maximum and minimum numbers of players of each role (Fig. 5B3).
- *Solution operations.* Users can click the generate team button to obtain the optimal lineup under the current input of tactics and players. In addition, they can delete the current lineup with the clear button and add the current lineup to the lineup view with the add button.
- *Switching player roles.* Users can click the switch button of player roles to switch the player role that occurs in the candidate player list or the explanation component.
- *Adjusting players.* Users can drag unselected players in the candidate player list to the icons in the lineup editor board to replace players according to their own opinions.

5.4 Lineup View

With multiple generated solutions in the player view, users can further identify the most suitable lineup in the lineup view with visual comparison. The lineup view includes a candidate lineup list (Fig. 5C1) and lineup thumbnails (Fig. 5C2) for effective decision-making (**L1**, **L2**). In this view, users can sort lineups in the candidate lineup list to decide on the most suitable lineup, and select lineups to obtain the details from lineup thumbnails for further comparison.

Candidate lineup list. The candidate lineup list records the solutions and provides a flexible comparison according to different criteria to choose the best lineup (Fig. 5C1). Each row of the list contains four components: the predicted match result, the confidence score, the predicted statistic indicators, and the formation icon [7] (Fig. 5C4). We use bar charts to present the lineup metrics to facilitate comparison.

The predicted match result is the probability of winning/drawing/losing the match. The confidence score is defined as the normalized average appearance time in the historical data of all players in the lineup, which could reflect the reliability of the lineup. The predicted statistic indicators include offensive indicators (i.e., passing number and the success rate of passing), scoring indicators (i.e., shot number and goal number), and defensive indicators (i.e., interception number and tackle number). Both the match result and the statistical indicators are predicted by the features of all players in the lineup with a random forest approach.

Lineup thumbnails. We design lineup thumbnails to present the details of candidate lineups (Fig. 5C2). When users select two lineups that need to be compared from the candidate lineup list, the thumbnails of the two lineups will be presented on the right of the list. The player icons in the lineup thumbnails are placed according to the player positions in the team formation. We add an arc around each player icon to encode the confidence score of the player. For convenient comparison among the lineups, those players who are not picked by all selected lineups are highlighted with black borders in the lineup thumbnails.

Interaction tools. Users can interact with the lineup view in various ways. The follow are some commonly used tools provided by the system.

- *Sorting lineups.* Users can sort lineups by each metric for effective multi-criteria decision-making.
- *Switching indicators.* Users can switch between different groups of statistic indicators to evaluate whether the lineup meets their expectations (Fig. 5C3).
- *Unfolding details.* Users can click lineups in the candidate lineup list to show the details with lineup thumbnails.

6 System Evaluation

In this section, we describe two case studies on the use of our system in selecting and understanding lineups. After the case studies, we also summarize the feedback of the experts.

6.1 Case Studies

We invited two experts to conduct two case studies on the use of our system. Our first expert is a senior analyst in sports science (EA), and our second expert is a senior soccer coach (EB). The cases used are from the soccer event dataset in our analysis with the matches from the five top European leagues in the 2017/18 season [56]. At the beginning of each case study, we introduced the design and interaction tools of our system. After the experts were familiar with our system, they tried to select and analyze the lineup of a team of interest. We recorded the analytical process and comments of each expert. The findings and the activities of the two case studies are summarized as follows.

6.1.1 Lineup Explanation: Why a Player is Selected by the Automatically Generated Lineup?

This insight was gained from the first case study conducted by EA. This case study is about exploring proper lineups for Barcelona when facing Real Madrid. These two teams are top teams in Spanish *La Liga*, and their competitions get significant attention among soccer analysts. As mentioned by EA, in the 2017/18 season, Barcelona won the championship of *La Liga* and recorded one win (0:3) and one draw (2:2) in matches against Real Madrid. Therefore, EA was interested in generating lineups under such a scenario.

Before exploring the lineups, EA selected Barcelona as the target team and Real Madrid as the opponent team, and investigated the tactics of Barcelona when facing Real Madrid in the tactic view. EA first focused on the defensive tactics of Barcelona and the offensive tactics of Real Madrid. In the distribution bar charts, EA filtered the offensive tactics of Real Madrid with passing because those tactics were used most frequently and the success rate is relatively high (Fig. 5A1). EA also filtered the defensive tactics of Barcelona with tackle and interception for the same reason (Fig. 5A2). After filtering tactic categories, EA switched to detailed tactics of both teams. Specifically, EA sorted the tactics by success rate and filtered out the tactics whose usage rates are less than 5% to examine important tactics. As for Real Madrid, EA noticed that the passing tactics with high success rates are mainly on the left-wing (Fig. 5A3). Thus, EA selected the first offensive tactic in the left list to filter those defensive tactics of Barcelona that are on the same spatial regions (Fig. 8A1). Thereafter, EA selected the defensive tactic of Barcelona with the highest success rate and assigned it to defenders and midfielders based on the tactical preference for defense (Fig. 8A2). Similarly, when specifying the offensive tactics of Barcelona, EA filtered the offensive tactics of Barcelona by the least effective defensive tactic of Real Madrid, the central-backfield defense (Fig. 8A3). After the filtering, EA assigned the most successful offensive tactic in the right list, mid-attack, to the forwards of Barcelona (Fig. 8A4).

After the tactics were specified, EA moved to the player view to generate and explore suitable lineups for Barcelona. EA first explored core players who needed to be included in the lineup. In the candidate player list, EA noticed that the learned weights of individual performance value and offensive interaction value were larger than defensive interaction value, meaning that these two features contribute most to the match results (Fig. 5B4). Thus, EA chose L. Messi as the core player in the lineup because of his excellent individual performance and offensive interactions with his teammates. Then, EA obtained the optimal lineup provided by the model and found that most players in the lineup met the expectations, including powerful forwards such as L. Suárez and outstanding midfielders like Iniesta (Fig. 8B).

Based on the lineup result, EA moved to the explanation component to explain why the players were selected by the model. EA mainly focused on the defenders because of the tactical preference for defense. On the defensive aspect, EA noticed that Piqué, Sergi Roberto, S. Umtiti, and Jordi Alba were ranked in the top four of all the defenders (Fig. 8B1). To be mentioned, Piqué, Sergi Roberto, and S. Umtiti were well-performed on defending the forwards and midfielders of the opponent team, and Jordi Alba was outstanding in defending the opposing midfielders (Fig. 8B3, B4). On the offensive aspect, EA found that Jordi Alba and Sergi Roberto could pass threatening balls to their teammates, especially forwards and midfielders (Fig. 7B2). Meanwhile, Piqué and S. Umtiti were able to pass valuable balls to other defenders to increase the goal chance. As for a detailed evaluation, EA unfolded all forwards and found that Jordi Alba and Sergi Roberto could create more opportunities for L. Messi and L. Suárez, who were also selected in the lineup and could rely on their excellent individual performances to score (Fig. 7C2). EA summarized that the four defenders were selected not only because they could prevent the attacking from the opponents, but also pass critical balls to their teammates to score a goal.

EA mentioned that the lineup recommended by the model had one different player compared with the lineup selected by the coach of Barcelona in the match that defeated Real Madrid. Specifically, the model selected the defender S. Umtiti, while the coach chose another defender T. Vermaelen. EA wondered why the lineup selection model chose S. Umtiti rather than T. Vermaelen. Thus, EA compared those two players in the explanation component. On the defensive aspect, EA found that S. Umtiti performed more effective defense on all roles of the opponent players than T. Vermaelen, especially on forwards (Fig. 8B4). This suggested that S. Umtiti, rather than T. Vermaelen, could steal threatening balls from the opposing forwards. On the offensive aspect, EA noticed that S. Umtiti created more valuable passes to midfielders, while T. Vermaelen interacted more closely with defenders (Fig. 8B5). This indicated that S. Umtiti was more likely to create chances through midfielders, while T. Vermaelen tended to pass to other defenders to start a build-up. EA further unfolded all midfielders and defenders for a detailed analysis. EA found that S. Umtiti passed more valuable balls to the skilled midfielder Iniesta, and T. Vermaelen interacted effectively with Sergi Roberto, the defender who was good at creating chances for forwards. EA concluded that S. Umtiti was selected by the model due to better performance on defense and more directly offensive interactions with midfielders.

Through this process, EA generated a lineup for Barcelona under a combination of tactics and explained why those players in the lineup were recommended by the system.

6.1.2 Adjusting Players: Which Player is More Suitable?

This insight was also gained from the first case study. After the process of the first insight, EA required to know whether the lineup generated in the system could be better than the lineup in the real match. Therefore, EA replaced S. Umtiti with T. Vermaelen in the generated lineup and added the two lineups to the lineup view for further comparison.

We refer to the lineup including S. Umtiti as L1 and the lineup including T. Vermaelen as L2. In the candidate lineup list, EA first examined the two lineups on the aspect of the confidence score and the predicted match result to identify the better lineup. EA found that the confidence score of L1 was higher than that of L2, meaning that the predicted performance of L1 is more robust than that of L2 (Fig. 8C2). Meanwhile, both the predicted winning rate and the chance of not losing the match of L1 were higher than those of L2 (Fig. 8C1). Based on the two aspects, EA summarized that L1 would be a better choice for the real match. Then, EA focused on the predicted statistical indicators to discover the details of the two lineups. EA switched among the three kinds of indicators and noticed that L1 would perform better on passing number and goal number, while other indicators were not different obviously. This confirmed the previously

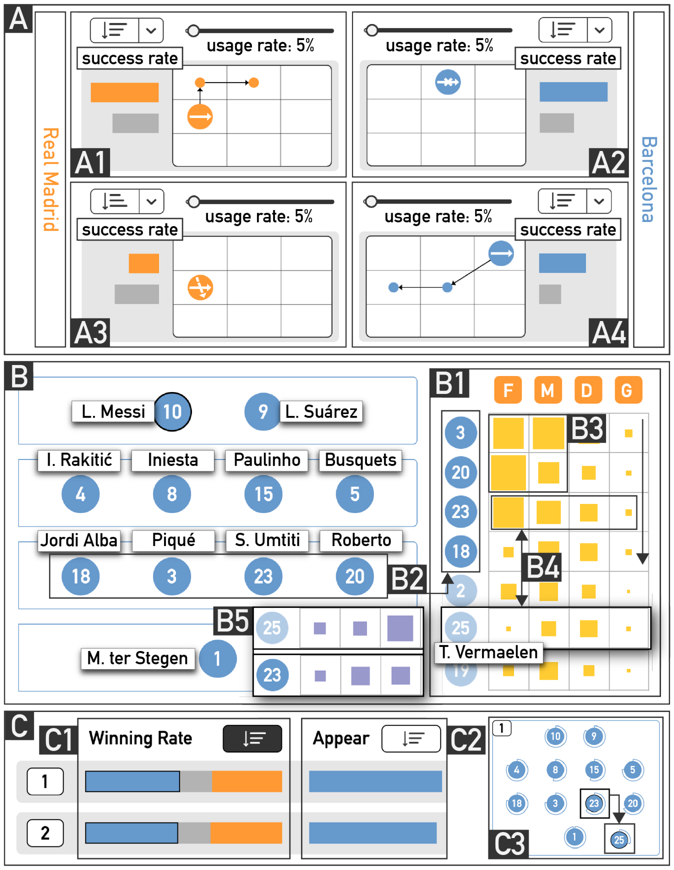


Fig. 8. Analysis pipeline in Case 1. (A) presents the tactic specifications. (B) shows the lineup generated for Barcelona when facing Real Madrid. (C) is the comparison between the generated and the adjusted lineup.

gained insight that the model selected S. Umtiti rather than T. Vermaelen because of his more effective interactions with excellent midfielders who could act as tactic organizers in matches, leading to more passes and a higher goal chance.

To inspect the detailed confidence score of each player, EA further clicked to display the thumbnails of the two lineups. EA found that the confidence scores of all the players who were selected by both lineups are more than a half (Fig. 8C3). This indicated that those players were performed stable and trusted to obtain extensive chances to appear in matches. As for the two different players, EA found that S. Umtiti attended more time in matches than T. Vermaelen. After comparing the two lineups, EA concluded that the model-generated lineup L1 would be more appropriate for the match with Real Madrid than the coach-selected lineup L2.

Through this process, EA adjusted the generated lineup from the perspective of coaches and discovered the differences between the expected performances of the two lineups. Such discoveries indicate that our system can provide more effective lineups for real matches.

6.1.3 Lineup Comparison: Which Lineup is the Best?

This insight was gained from the second case study conducted by EB. This case study is about choosing a lineup for Manchester United when competing with Manchester City. These two teams were the top two in the 2017/18 season of the *English Premier League*. As mentioned by EB, Manchester United got one lose (1:2) and one win (3:2) when facing Manchester City in the season. EB was interested in improving the lineup of Manchester United with our system.

EB began with the exploration from strategy specification in the tactic view. Firstly, EB filtered the passing tactics of Manchester City and the interception tactics of Manchester United, selected the most effective passing tactic of Manchester City, and assigned the most successful interception tactic on the same spatial region to the defenders (Fig. 9A1, A3). Afterwards, EB focused on the interception tactics of Manchester City and the passing tactics of Manchester United, selected the least used interception tactic of Manchester City, and assigned the corresponding passing tactic to the midfielders and forwards due to the tactical preference for offense (Fig. 9A2, A4). Then, EB turned to the player view to get lineups for Manchester United. EB chose P. Pogba as the core player due to his outstanding offensive interaction with his teammates (Fig. 9B), and obtained three lineups under the current tactic selection. The first lineup was obtained directly without additional operations. The second lineup was constrained by a 4-3-3 formation, which was often used by Manchester United through the season. The third lineup was adjusted by EB based on the second lineup. More specifically, EB replaced an offensive midfielder, J. Lingard (Fig. 9B2), with H. Mkhitaryan (Fig. 9B1), another offensive midfielder with better offensive interactions with the teammates, to find out whether a lineup with cooperative but less skilled players for Manchester United would perform better. EB also generated another lineup with a different tactic selection by replacing the tactic for the midfielders with the most successful free-kick tactic as it was another essential attacking method for Manchester United (Fig. 9C).

EB added those four lineups (referred as L1, L2, L3, and L4) into the lineup view to examine the overall performances with multi-criteria comparison among different indicators (Fig. 9D). EB first clicked the table headers of the winning rate and the appearance time to identify the most appropriate lineup. EB found that L1 and L2 ranked the top two under the winning rate (Fig. 9D1, D2), and all the three lineups under the first tactic selection were higher than L4 under the appearance time (Fig. 9D3). Such comparison results suggested that the first tactic selection would be more appropriate for matches against Manchester City. Besides, EB mentioned that L2 was better than L3 on both the winning rate and appearance time (Fig. 9D3), indicating that the lineup with players only outstanding in passing threatening balls would perform worse than balanced lineups under the first tactic specification. Furthermore, EB noticed that L1 was the best in winning the match but was ranked third in robustness (Fig. 9D1). Meanwhile, the expected performance of L2 was the most robust and ranked second in the winning rate (Fig. 9D2). It meant that coaches are required to make a trade-off between the two lineups. Thus, EB turned on the thumbnails of L1 and L2 for a detailed comparison.

Through the lineup thumbnails, EB found that L1 selected the offensive midfielder M. Fellaini, who performed well but seldom appeared in the historical matches, and L2 chose the more experienced defender P. Jones (Fig. 9E1, E2). As for the statistical indicators, EB found that both the two scoring indicators of L1 were higher than L2 (Fig. 9E5). This confirmed the hypothesis of EB that L1 was a typical offensive formation 3-4-3, and the offensive midfielder M.

Fellaini was well-performed in attacking than the defender P. Jones. EB further switched to the other two groups of indicators and found that L1 would create slightly fewer passes (Fig. 9D) and more tackles than L2 (Fig. 9E6). It also corresponded to the knowledge of EB that L1 would be more suitable for the high-press strategy, while L2 could be adapted for the build-up strategy. EB concluded that if Manchester United took offensive strategies, L1 could be the most suitable lineup, and L2 would be the most proper lineup if more stable strategies were employed.

Through this process, EB obtained four different lineups for Manchester United with different tactic combinations when facing Manchester City. EB also gained insights about comparing those lineups and choosing the most suitable one under different scenarios.

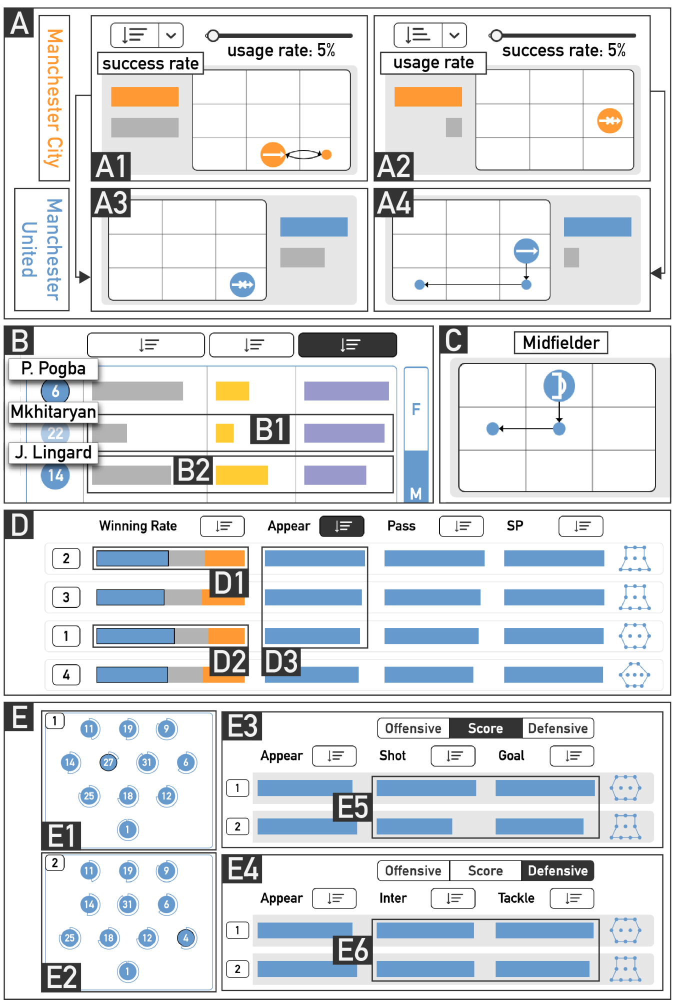


Fig. 9. Analysis pipeline of Case 2. (A) presents the tactic specifications. (B) and (C) illustrate the lineup adjustments. (D) shows the four lineups generated for Manchester United. (E) presents the multi-criteria comparison of lineups.

6.2 Expert Interview

We interviewed the two experts who conducted the case studies (EA and EB) and the other two new experts (EC and ED) to evaluate the usefulness and effectiveness of our system. Both the two new experts are doctoral students in physical education and are experienced in soccer data analysis. We interviewed those experts respectively and collected their feedback during the interview.

Procedure. The interview procedure for the two new experts (EC and ED) contains four steps. First, we briefly introduced the metrics and the model used in our soccer lineup selection system. Then, we presented the visual design and interactions to the experts through a usage scenario. We also let the experts freely explore the system and answered their questions to help them become familiar with the system. Afterward, the experts were required to select the most appropriate lineup for Barcelona when facing Real Madrid. We recorded their analysis procedures during this step. Finally, we asked the experts whether each design requirement (i.e., T1-T2, P1-P2, L1-L2) has been solved by the visual design, whether there are potential issues in the current design, and whether the system is sufficient to support a more effective and efficient lineup selection process. After EA and EB finished the case study, we also asked those questions to gain their feedback.

System usability. Generally, all the experts were satisfied with our system and felt that it could help soccer coaches select the lineup effectively. Compared with directly selecting players by experience, the coaches could specify their tactical preferences based on the tactic detection results and generate lineups interactively with our system. EA commented that *"Coaches usually select lineups by their own opinions on player performance, which is laborious and subjective. Such a system could provide them an efficient method to select lineups based on historical data and evaluate them by expected performance"*. The experts also thought highly of our lineup selection model because it considers the relationships between players and could provide more reasonable lineup results. EB indicated that *"Most of the players selected by the system matched my expectations. Besides, the system could recommend new lineups I have not thought of previously, which inspired me to explore whether a known lineup could be improved through lineup adjustments and the prediction of lineup performance"*. They also mentioned that with the system, coaches could understand why the players were selected by the model and decide whether apply the lineup in real matches.

Visual design. All the experts agreed that each design requirement has been fulfilled by our visual design. Specifically, ED appreciated the visual design for tactic specification, *"... the confrontation tactic list is helpful in discovering essential tactics, and the presentation of tactics are effective for identifying the tactic at a glance"*. EB and EC liked the candidate player list and mentioned that it could be particularly useful in the scenario of selecting the lineup for a real match. *"The comparison and ranking would be effective when selecting the core players"* (EC). *"Injuries and match bans commonly occur in lineup selection. I did not use the function of removing players in the case study because of the lack of data. I can remove those players when selecting lineups in real matches"* (EB). As for the explanation component, EA was impressed by the matrix-based design, *"It is convenient to find player pairs with successful interactions by the matrix"*. ED also favored the multi-criteria comparison of the lineups, *"... it can help me make a trade-off among multiple lineups and quickly find the one that I desired"*.

Suggestions. The experts had some suggestions for improvement. EA focused on the lineup selection model and hoped to add movement data of players without the ball for a more comprehensive evaluation of player interactions. EB and EC considered the interactions in the player candidate list and believed that a search bar could be more effective for finding players who cannot appear in the match. ED concentrated on the system design and suggested providing a record list for specifications and constraints for all lineups.

7 Discussion

Significance. Lineup selection plays an important role in team sports, such as basketball and soccer. Existing work on automatic soccer lineup selection is weak in incorporating the inputs from coaches and supporting interactive exploration of lineup possibilities. In this research, we develop an approach to address these challenges through modeling player performances and developing an interactive visual analytics system for lineup generation, explanation, and comparison. Our approach allows coaches to steer the selection and choose the most appropriate lineup based on their own opinions and the opponent information.

Our approach can be generalized to other team sports with the requirements on personal and team skills similar to soccer, such as basketball, ice hockey, and baseball. Our model is intrinsically composed of the summarization of individual performance of players and player interactions in matches, and can be converted for other team sports by incorporating appropriate evaluation metrics of individual performance and player interactions. Our visual design can also be easily modified based on the characteristics of sports.

Lessons learned. We have learned two lessons through the research. The first lesson is from the evaluation of defensive interactions. Defending in soccer matches is difficult to evaluate because it cannot directly create goals. During the meetings, the experts indicated that the defensive interactions mainly contribute to preventing the goal of the opponent team and increasing the goal chance of the team. In this way, the defensive interaction between two players from different teams can be evaluated through the change of goal probability of the two teams. The second lesson is from the visualization of offensive tactics. Compared with placing all phases in the tactic on the soccer pitch (the left part of Fig. 6A), an aggregated sequence of the tactic (the right part of Fig. 6A) could help users clearly identify how the ball was passed through the pitch. We further divide the pitch into nine regions to match offensive tactics with defensive tactics (Fig. 6B). The case study has shown that the aggregated sequence and pitch division are effective in the exploration of confrontation tactics. Such designs could also be generalized to other team sports such as basketball and ice hockey.

Limitations. There are some limitations in this research. First, our model does not include special factors that may influence a lineup, such as the physical and mental conditions of players, due to lack of relevant data. These factors should be considered when coaches need to select more reliable lineups for the match. We plan to collect such data and integrate them into the measurement of player teamwork to improve the reliability of our model. Second, the player role constraints in our model are limited to the four basic roles (i.e., goalkeeper, defender, midfielder, forward). The preferred side indicated by the detailed role (e.g., left-winger, center forward) is not considered. These fine-grained constraints should be considered if more detailed specification of lineups is required. We will develop more sophisticated models to improve the player role constraints to deal with the preferred side as the future work.

8 Conclusion

In this paper, we present a method for visual analytics of soccer lineup selection. Collaborating with the domain experts, we characterize the problem and propose a new teamwork-based model to integrate the factors preferred by coaches (e.g., spatial regions, opponent information) into lineup selection. Based on the model, we develop a web-based visual analytics system, Team-Builder, to allow coaches to interactively generate, explain, and compare lineups produced by an automatic model, and integrate their own opinions into lineup selection.

In the future, our work can be improved in several ways. One direction is to enhance our model by considering more data types (e.g., movement of players without the ball, interactions among multiple players in a period of time) so that coaches can evaluate the performances of players and their interactions more comprehensively. With these data, our lineup metrics can also be integrated into other analysis tasks such as evaluating the change of team formation. Another direction is to extend our work to other team sports such as basketball, ice hockey, and baseball by adapting our models, algorithms, and visualization tools based on the characteristics of individual sports. Besides, we also plan to improve the system usability based on the long-term expert feedback for the real lineup selection scenario.

Acknowledgments

We would like to thank Jiahui Chen and Rui Yang from the State Key Lab of CAD&CG, Zhejiang University, for their contribution to system development. The work was supported by NSFC (62072400) and the Collaborative Innovation Center of Artificial Intelligence by MOE and Zhejiang Provincial Government (ZJU). This work was also partially funded by the Zhejiang Lab (2021KE0AC02).

References

[1] B. H. Boon and G. Sierksma, "Team Formation: Matching Quality Supply and Quality Demand," *European Journal of Operational Research*, vol. 148, no. 2, pp. 277–292, 2003.

[2] H. Ruiz, P. Power, X. Wei, and P. Lucey, ""The Leicester City Fairytale?": Utilizing New Soccer Analytics Tools to Compare Performance in the 15/16 & 16/17 EPL Seasons," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017, pp. 1991–2000.

[3] UEFA, "EURO 2020 Technical Report," https://uefatechnicalreports.com/uefa-euro-2020, 2021.

[4] E. Ozceylan, "A Mathematical Model Using AHP Priorities for Soccer Player Selection: A Case Study," *South African Journal of Industrial Engineering*, vol. 27, no. 2, pp. 190–205, 2016.

[5] S. A. F. Salles, H. R. M. da Hora, M. E. Júnior, A. C. de Souza Gomes dos Santos, and A. Shimoya, "Operations Research Contributions for Football Teams Formation: A Systematic Review," *Pesquisa Operacional*, vol. 39, no. 2, pp. 277–293, 2019.

[6] G. L. Andrienko, N. V. 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*, vol. 27, no. 4, pp. 2280–2297, 2019.

[7] 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.

[8] 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, 2021.

[9] A. Cao, X. Xie, J. Lan, H. Lu, X. Hou, J. Wang, H. Zhang, D. Liu, and Y. Wu, "MIG-Viewer: Visual Analytics of Soccer Player Migration," *Visual Informatics*, vol. 5, no. 3, pp. 102–113, 2021.

[10] D. McEwan and M. R. Beauchamp, "Teamwork in Sport: A Theoretical and Integrative Review," *International Review of Sport and Exercise Psychology*, vol. 7, no. 1, pp. 229–250, 2014.

[11] D. W. Eccles and G. Tenenbaum, "Why an Expert Team is More Than a Team of Experts: A Social-Cognitive Conceptualization of Team Coordination and Communication in Sport," *Journal of Sport and Exercise Psychology*, vol. 26, no. 4, pp. 542–560, 2004.

[12] D. W. Eccles, "The Coordination of Labour in Sports Teams," *International Review of Sport and Exercise Psychology*, vol. 3, no. 2, pp. 154–170, 2010.

[13] J. Kuehn, "Accounting for Complementary Skill Sets When Evaluating NBA Players' Values to a Specific Team," in *2016 MIT Sloan Sports Analytics Conference*, vol. 6, 2016.

[14] A. Z. Maymin, P. Z. Maymin, and E. Shen, "NBA Chemistry: Positive and Negative Synergies in Basketball," *International Journal of Computer Science in Sport*, vol. 12, no. 2, pp. 4–23, 2013.

[15] R. Beal, N. Changder, T. J. Norman, and S. D. Ramchurn, "Learning the Value of Teamwork to Form Efficient Teams," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 5, 2020, pp. 7063–7070.

[16] L. Bransen and J. Van Haaren, "Player Chemistry: Striving for a Perfectly Balanced Soccer Team," *arXiv Preprint*, vol. arXiv, p. 2003.01712, 2020.

[17] J. Fernandez-Navarro, L. Fradua, A. Zubillaga, P. R. Ford, and A. P. McRobert, "Attacking and Defensive Styles of Play in Soccer: Analysis of Spanish and English Elite Teams," *Journal of Sports Science*, vol. 34, no. 24, pp. 2195–2204, 2016.

[18] C. Perin, R. Vuillemot, C. D. Stolper, J. T. Stasko, J. Wood, and S. Carpendale, "State of the Art of Sports Data Visualization," *Computer Graphics Forum*, vol. 37, no. 3, pp. 663–686, 2018.

[19] 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.

[20] A. Franks, A. Miller, L. Bornn, and K. Goldsberry, "Counterpoints: Advanced Defensive Metrics for NBA Basketball," in *MIT Sloan Sports Analytics Conference*, 2015.

[21] 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.

[22] W. Wang, J. Zhang, X. Yuan, and S. Liu, "MatchOrchestra: A Generalized Visual Analytics for Competitive Team Sports," *Journal of Visualization*, vol. 19, no. 3, pp. 515–528, 2016.

[23] J. P. Ono, C. A. Dietrich, and C. T. Silva, "Baseball Timeline: Summarizing Baseball Plays Into a Static Visualization," *Computer Graphics Forum*, vol. 37, no. 3, pp. 491–501, 2018.

[24] M. Lage, J. P. Ono, D. Cervone, J. Chiang, C. A. 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.

[25] C. A. Dietrich, D. Koop, H. T. Vo, and C. T. Silva, "Baseball4D: A Tool for Baseball Game Reconstruction and Visualization," in *Proceedings of IEEE Conference on Visual Analytics Science and Technology*, 2014, pp. 23–32.

[26] B. Moon and R. Brath, "Bloomberg Sports Visualization for Pitch Analysis," in *Workshop on Sports Data Visualization*, 2013.

[27] 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.

[28] 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, 2021.

[29] J. Wu, D. Liu, Z. Guo, Q. Xu, and Y. Wu, "TacticFlow: Visual Analytics of Ever-Changing Tactics in Racket Sports," *IEEE Transactions on Visualization and Computer Graphics*, vol. 28, no. 1, pp. 835–845, 2022.

[30] J. Lan, J. Wang, X. Shu, Z. Zhou, H. Zhang, and Y. Wu, "Rally-Comparator: Visual Comparison of the Multivariate and Spatial Stroke Sequence in Table Tennis Rally," *Journal of Visualization*, vol. 25, no. 1, pp. 143–158, 2022.

[31] P. A. Legg, D. H. S. 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.

[32] 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.

[33] Y. Ishikawa and I. Fujishiro, "TideGrapher: Visual Analytics of Tactical Situations for Rugby Matches," *Visual Informatics*, vol. 2, no. 1, pp. 60–70, 2018.

[34] C. Perin, R. Vuillemot, and J.-D. Fekete, "A table!: Improving Temporal Navigation in Soccer Ranking Tables," in *CHI Conference on Human Factors in Computing Systems*, 2014, pp. 887–896.

[35] C. Perin, J. Boy, and F. Vernier, "Using Gap Charts to Visualize the Temporal Evolution of Ranks and Scores," *IEEE Computer Graphics and Applications*, vol. 36, no. 5, pp. 38–49, 2016.

[36] A. Rusu, D. Stoica, and E. Burns, "Analyzing Soccer Goalkeeper Performance Using a Metaphor-Based Visualization," in *2011 15th International Conference on Information Visualisation*, 2011, pp. 194–199.

[37] A. Rusu, D. Stoica, E. Burns, B. Hample, K. McGarry, and R. Russell, "Dynamic Visualizations for Soccer Statistical Analysis," in *2010 14th International Conference on Information Visualisation*, 2010, pp. 207–212.

[38] 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.

[39] J. L. S. Malqui, N. M. L. Romero, R. Garcia, H. Alemdar, and J. L. Comba, "How Do Soccer Teams Coordinate Consecutive Passes? A Visual Analytics System for Analysing the Complexity of Passing Sequences Using Soccer Flow Motifs," *Computers and Graphics*, vol. 84, pp. 122–133, 2019.

[40] 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.

[41] M. Stein, H. Janetzko, A. Lamprecht, T. Breitkreutz, P. Zimmermann, B. Goldlücke, T. Schreck, G. L. 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.

[42] L. Shao, D. Sacha, B. Neldner, M. Stein, and T. Schreck, "Visual-Interactive Search for Soccer Trajectories to Identify Interesting Game Situations," *Electronic Imaging*, no. 1, pp. 1–10, 2016.

[43] D. Sacha, F. Al-Masoudi, 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.

[44] G. L. Andrienko, N. V. Andrienko, G. Budziak, J. Dykes, G. Fuchs, T. von Landesberger, and H. Weber, "Visual Analysis of Pressure in Football," *Data Mining and Knowledge Discovery*, vol. 31, no. 6, pp. 1793–1839, 2017.

[45] H. Janetzko, D. Sacha, M. Stein, T. Schreck, D. A. Keim, and O. Deussen, "Feature-Driven Visual Analytics of Soccer Data," in *2014 IEEE Conference on Visual Analytics Science and Technology (VAST)*, 2014, pp. 13–22.

[46] V. Machado, R. Leite, F. Moura, S. Cunha, F. Sadlo, and J. L. Comba, "Visual Soccer Match Analysis Using Spatiotemporal Positions of Players," *Computers and Graphics*, vol. 68, pp. 84–95, 2017.

[47] M. Stein, D. Seebacher, R. Marcelino, T. Schreck, M. Grossniklaus, D. A. Keim, and H. Janetzko, "Where to Go: Computational and Visual What-if Analyses in Soccer," *Journal of Sports Sciences*, vol. 37, no. 24, pp. 2774–2782, 2019.

[48] A. B. Santos, R. Theron, A. Losada, J. E. Sampaio, and C. Lago-Peñas, "Data-Driven Visual Performance Analysis in Soccer: An Exploratory Prototype," *Frontiers in Psychology*, vol. 9, p. 2416, 2018.

[49] L. Shaw and M. Glickman, "Dynamic Analysis of Team Strategy in Professional Football," *Barça Sports Analytics Summit*, vol. 13, 2019.

[50] A. Bialkowski, P. Lucey, P. Carr, Y. Yue, S. Sridharan, and I. Matthews, "Large-Scale Analysis of Soccer Matches Using

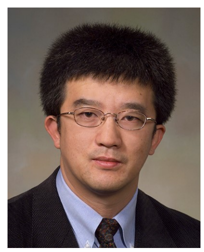
Spatiotemporal Tracking Data," in *2014 IEEE International Conference on Data Mining*, 2014, pp. 725–730.
[51] J. Perl and D. Memmert, "Soccer Analyses by Means of Artificial Neural Networks, Automatic Pass Recognition and Voronoi-Cells: An Approach of Measuring Tactical Success," in *Proceedings of the 10th International Symposium on Computer Science in Sports*, 2016, pp. 77–84.
[52] M. Tavana, F. Azizi, F. Azizi, and M. Behzadian, "A Fuzzy Inference System with Application to Player Selection and Team Formation in Multi-player Sports," *Sport Management Review*, vol. 16, no. 1, pp. 97–110, 2013.
[53] J. M. Merigó and A. M. Gil-Lafuente, "Decision-making in Sport Management based on the OWA Operator," *Expert Systems with Applications*, vol. 38, no. 8, pp. 10 408–10 413, 2011.
[54] R. Al-Shboul, T. Syed, J. Memon, and F. Khan, "Automated Player Selection for Sports Team using Competitive Neural Networks," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 8, pp. 457–460, 2017.
[55] T. Decroos, L. Bransen, J. Van Haaren, and J. Davis, "Actions Speak Louder than Goals: Valuing Player Actions in Soccer," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019, pp. 1851–1861.
[56] L. Pappalardo, P. Cintia, A. Rossi, E. Massucco, P. Ferragina, D. Pedreschi, and F. Giannotti, "A Public Data Set of Spatio-temporal Match Events in Soccer Competitions," *Scientific Data*, vol. 6, no. 1, pp. 1–15, 2019.
[57] T. Decroos, J. Van Haaren, and J. Davis, "Automatic Discovery of Tactics in Spatio-Temporal Soccer Match Data," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2018, pp. 223–232.
[58] P. Power, H. Ruiz, X. Wei, and P. Lucey, ""Not All Passes Are Created Equal": Objectively Measuring the Risk and Reward of Passes in Soccer from Tracking Data," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017, pp. 1605–1613.
[59] S. Latif and F. Beck, "Interactive Map Reports Summarizing Bivariate Geographic Data," *Visual Informatics*, vol. 3, no. 1, pp. 27–37, 2019.
[60] W. Zeng, C.-W. Fu, S. M. Arisona, S. Schubiger, R. Burkhard, and K.-L. Ma, "A Visual Analytics Design for Studying Rhythm Patterns from Human Daily Movement Data," *Visual Informatics*, vol. 1, no. 2, pp. 81–91, 2017.
[61] B. Ni, Q. Shen, J. Xu, and H. Qu, "Spatio-Temporal Flow Maps for Visualizing Movement and Contact Patterns," *Visual Informatics*, vol. 1, no. 1, pp. 57–64, 2017.

Anqi Cao is currently a Ph.D. student of computer science at the State Key Lab of CAD&CG, Zhejiang University. Her research interest includes sports data visualization and visual analytics. She received her Bachelor's Degree in computer science from Nankai University in 2019.

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/.

Xiao Xie is an Assistant Professor 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.

Hongyu Chen is currently a senior majoring in Computer Science and Technology in Zhejiang University. Her research interest includes sports data visualization and visual analytics.



Xiaolong (Luke) Zhang is an Associate Professor at the College of Information Sciences and Technology, Pennsylvania State University. His research interests is in using advanced interactive technologies like multi-scale, virtual environments, and collaboration to improve the understanding of large information structures and the design of interactive tools to support visually-guided navigation in large information space. Currently, he is working on issues concerning multi-facet visualization of information structures and multi-scale navigation. He received his Ph.D. degree in School of Information from University of Michigan.

Hui Zhang graduated from Zhejiang Normal University with bachelor's degree, Beijing Sport University with Master's and University of Potsdam in Germany for Ph.D. degree. 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 a 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 the Department of Sports Science, College of Education, Zhejiang University as a Professor.

Yingcai Wu is a Professor at the State Key Lab of CAD&CG, Zhejiang University. His main research interests are in visual analytics and information visualization, with focuses on sports analytics and big data intelligence. 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 in 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.