FIFA19 Data Visualization Tool

Tairen Piao*
Seoul National University

Huiwen Xu†
Seoul National University

ABSTRACT

Association football is the most popular sport in the world and that made football a considerable entertainment business operating worldwide. People want easily and quickly to get information about players who they are interested in. Therefore our target users are general football fans, who would like to know their favorite players' profiles and personal skills, and rankings according to some attributes. In this project, we implement an interactive tool that visualizes football players' skills and latent information by utilizing FIFA data. Our task consists of two parts. First, we extract latent information, which will help for analyzing similarity among football superstars. Second, we visualize the information for target users. As a result, target users can easily dig out the significant information from the visualization charts.

Index Terms: Visualization tool, FIFA19, HCI

1 INTRODUCTION

Football is a massive entertainment business operating worldwide. It is estimated that more than half of the world's population consider themselves to be association football fans. People want easily and quickly to get information about players whom they are interested in. However, different target users have different needs. Our target users are general football fans, who would like to know their favorite players' profiles and personal skills, and rankings according to some attributes.

To tackle this issue, we provide several figures to meet users' satisfaction, e.g., hexagon charts for individual skill, bar charts for top 20 players with their position, and some specific attributes. We add some interactions for users, which are more effective in visualizing our charts. We use FIFA19 datasets, which have been used for analysis provides statistics of about 18000 players on over 90 different attributes. In this project, we have two main ideas: first, extract latent information, which will help for analyzing similarity among football superstars; second, visualize the data for target users. As a result, target users can easily dig out the significant information from the visualization charts.

Section 2 describes domain situations, including target users, goals, and tasks. Dataset and target abstraction is presented in Section 3. Our visualization designs and implementations are described in Section 4, 5, respectively followed by illustrations. Section 6 shows the evaluation of our method. After discussing in Section 7, we conclude in Section 8.

2 DOMAIN SITUATION

2.1 Target Users

Our main target user is football fans both fans and momentary fans. A "football fan" by cultural definition is a generic sports follower who doesn't really follow a particular club or league, but takes interest at times when major competitions or games like the World Cup finals, Champions League finals, the El Classico and Cup

---

*e-mail: piaotairen@snu.ac.kr
†e-mail: xuhuiwen33@snu.ac.kr

finals occur. They are momentary fans, who like to watch the game but can't really tell you if Neymar played in the last World Cup or not.

2.2 Motivations and Goals

There are many visualizations about FIFA now. Our motivation is to provide football fans with a friendly and concise visualization. We hope that football fans will find the information they want to see. And we have several goals.The first goal is to give them some direct and clear visualization about some necessary FIFA information make them know more about football players, football clubs and other information.The second goal is to give the football fans some detailed information that they may interest in. Charts will provide more professional information with some simple data analysis skills. Finally, we will give a chart with a selection function. The third goal is to provide professional data or information to the target user. We plan to find some hidden information through data mining skills and algorithms.

Overall, Our main goal is to provide football fans with some information about the latest FIFA19 of clubs and players through visualization, which is interactive, and to provide users with diverse information through different charts, and also perform some data analysis to users Provide some information hidden in the data.

2.3 Demand Analysis

Our target user is football fans. The charts that football fans want to see should have the following characteristics.

- **To be interactive.** Give users a choice of features that both understand what they want to know.

- **Friendly interface.** A friendly interface is necessary for every system or tool.

- **Efficient.** Users want faster and more efficient access to information, and the purpose of converting '.csv' tables into visualizations lies in this.

- **Avoid chartjunk.** User would not want to use this tool anymore if there are too much useless information

3 DATA AND TASK ABSTRACTION

3.1 Data

The data is scraped from the sofifa website[1] using a python crawling script. The website contains the data from the EA Sports' game FIFA and gets updated regularly with the release of new versions of the game. Data developed by Electronic Arts for the latest edition of their FIFA game franchise. Since 1995 the FIFA Soccer games provide an extensive and coherent scout of players worldwide.

The FIFA19 dataset that has been used for this analysis provides statistics of about 18000 players on over 90 different attributes as shown in 1. These attributes are optimal indicators to determine the performance of a player at a particular playing position. For each attribute, we have an integer from 0 to 100 that measures how good a player is at that attribute. Examples of attributes are dribbling, aggression, vision, marking and ball control. Observe that it seems

---

[1]https://sofifa.com

to be unfeasible to accurately characterize players in these attributes automatically. Thus, all of those are gathered and curated by the company whose job is to bring the gameplay closer to reality as possible, hence preserving coherence and representativeness across the dataset.

3.2 Task Abstraction

There are three main task abstraction levels [2]. Choices at each level are independent

- High-level: Analyze with actions consume, and produce.

- Mid-level: Search with actions lookup, browse, locate and explore.

- Low-level: Query with actions identify, compare and summarize.

For task abstraction we did query and search. In low-level query we did compare, and in mid-level search we did browse and explore. For more specific we have three main tasks as follows. For the first task, we provide charts about the necessary information in FIFA. We give a radar chart and a line chart with five attributes, which represent the five central skill values of the player. When the user selects a player, the skill value will be displayed in the radar chart and the line chart. For the second task, we provide a bar chart with many interactions. Users can choose the skill value they want to pay attention to. The chart will reflect the player rankings under this skill value, and they can decide how many players to watch. For the third task, We provide an interactive line chart. The user selects two attributes that they want to view. The chart returns a correlation. The points in the chart represent players, and the lines represent the correlation.

4 VISUALIZATION DESIGN

Firstly, The main page is divided into two parts, the left side and the right side. The left side occupies 15% of the page, and the right side holds 85%. The left side is mainly for clubs, player visualization options, etc. The right is mainly for different charts. The head of the page is a propaganda image of FIFA19, which allows users to know that it is visual data about FIFA19 instantly. The right part of the main interface introduces our project. In the player selection bar on the left, we have added interactivity. When the mouse is moved over the player's name, the tool will return a visual feedback. The content is information about the player. The figure is shown in the Fig. 1.

(a) Select options
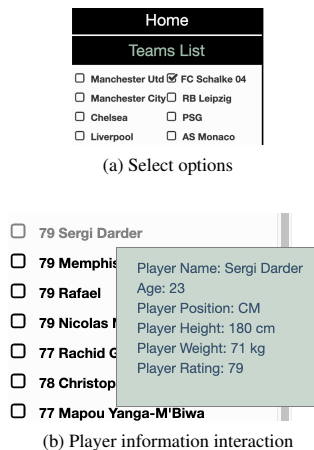

(b) Player information interaction

Figure 1: Selection bar and information interaction

For the radar chart, because each player's skill value can be summarized by six skill values, the radar chart has six angles, 0 is at the center and 100 is at the edge. The closer to the edge, the higher the skill value. On the right side of the radar chart, there is a line chart that expands it. You can make a horizontal comparison of the capability values, as shown in Fig.2.
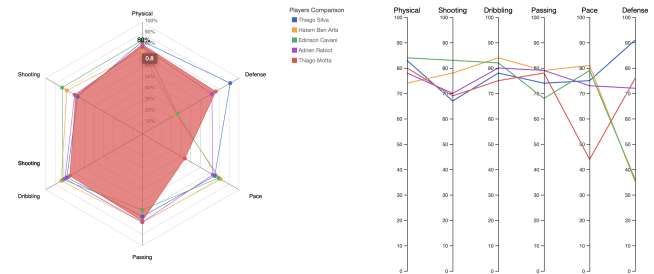

Figure 2: Radar chart and line chart of player skills

Below the radar chart and the line chart are another line chart and the top 10 players radar chart, respectively. There are multiple buttons above the line chart, which represent six skills values. After clicking the buttons, the corresponding line value will be expanded below the line chart. At the bottom, there is a radar chart of top 10 players. With the increase in the number of players selected, only one radar chart can not ultimately see the players with outstanding skill values so that users can see those players with the outstanding skill through these radar charts. The line chart and top 10 radar charts are shown in the Fig.3.
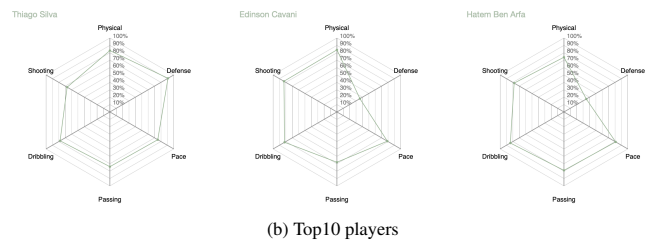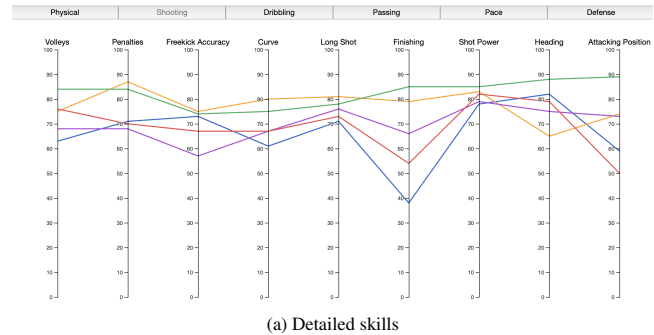

(a) Detailed skills

(b) Top10 players

Figure 3: Detailed line chart of one skill

For the bar chart, because the user is not only interested in the player's comprehensive skill value, but may be interested in a certain skill value, we added a button according to the affiliation relationship. When the user selects a different skill value, it will return a ranking of the bar chart And the number of players can be adjusted according to the Load more and Load less buttons.as shown in Fig.4.

The last chart is the correlation line chart. Firstly, two options are designed for users to choose the correlation they want to view. After selection, a line chart will appear below. There are different

Name	Age	Nationality	Club	Wage	Position	Potential	Stamina	Overall	Dribbling	...
L. Messi	31	Argentina	FC Barcelona	€565K	RF	94	72	94	97	...
Cristiano Ronaldo	33	Portugal	Juventus	€405K	ST	94	88	94	88	...
Neymar Jr	26	Spain	Paris Saint-Germain	€290K	LW	93	81	92	96	...
De Gea	27	Belguim	Manchester United	€260K	GK	93	43	91	18	...
K. De Bruyne	27	Belguim	Manchester City	€355K	RCM	92	90	91	86	...

Table 1: FIFA19 Complete Player Dataset

Figure 4: Bar chart of single attribute

color points in the chart which represent different values. In the right corner there is a bar represent the value density of a point, and the horizontal and vertical coordinates represent two attributes respectively. The line represents correlation, as shown in Fig.5.
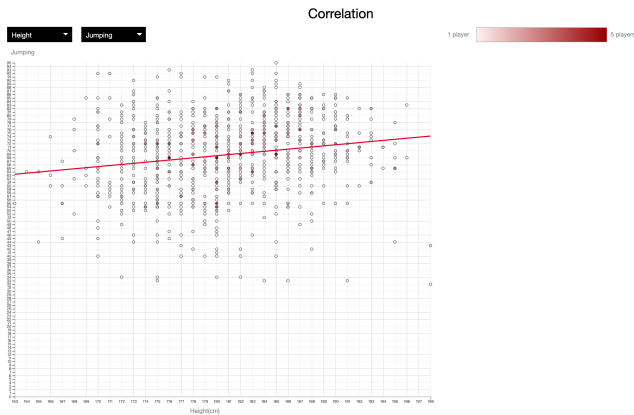

Figure 5: Correlation line chart

5 IMPLEMENTATION

5.1 Tool

We implemented most of the charts with d3 version 5,Front-end pages are implemented using HTML + CSS + Javascript. In order to make the elements in the page like <text>, <button> and <div> look more friendly, we use Bootstrap3 to beautify the HTML page.

5.2 Data Porcessing

We downloaded FIFA19 datasets in kaggle and processed the data using a python file. Firstly, We use the average method to complete the missing information in the data, and then We get a different .csv file with rows that we want. We use d3 to connected data in the main.js file and make calls.

5.3 Front-end page

First Create a project and make a new html file and introduce the css and js files that we need.Almost all layouts are implemented using <div> boxes, including element position, size and style,Finally, the required chart is introduced into the <div> block.

5.4 Charts

For the radar chart, First bind the data and use d3 to implement the chart. The edges are implemented with axises, and the values are implemented with lines. For interaction we use d3.transition to implement the filling background effect.

For the bar chart, first we bind the corresponding data according to the option selected by the option, and then uses d3 to implement a coordinate axis, the ordinate represents the name of the player, and the abscissa represents the player's skill value in the option.

For the line chart, first we bind the corresponding data, and we use d3 to implement the coordinate axis. The abscissa is the player's skill value, and the ordinate is the other skill value of the players, but a specific value of different players may have the same data, so as a result of multiple points in one position, we use d3 color to distinguish the number, the darker the color of the point represents the more significant the amount of data.

6 EVALUATION

The System Usability Scale (SUS) provides a "quick and dirty", reliable tool for measuring the usability. It consists of a 10 item questionnaire with five response options for respondents; from Strongly agree to Strongly disagree. Originally created by John Brooke in 1986, it allows you to evaluate a wide variety of products and services, including hardware, software, mobile devices, websites and applications [1]. We did some simple statistics and evaluation of our tool after the presentation. First, we evaluated the speed of the tool. Although the amount of data is large, our speed can respond on time. We evaluated the tool's visualization design. It can be seen that the radar chart is very suitable for showing the player's skill. With the effect of pop out, expanding the skill value to a line chart can make the player's skill value more precisely.

However, there is a duplication between our expansion skill line chart and the skill line chart. The user will be inexplicable. You need to confirm to know what this chart is. In the visualization of selecting players on the left, the selection is very convenient, and the options have visual feedback. At the same time, the font is beautified using Bootstrap, and the user experience is better. In the correlation line chart, we use dots with different colors to indicate the number. Users can quickly find and intuitively determine the correlation between two attributes. On the whole, our tool provides a variety of visual

Features	Detail skills
Physical	Stamina, Strength, Jumping
Shooting	Volleys, Penalties, Free-kick Accuracy, Curve, Long Shot, Finishing, Shot Power, Heading, Attacking Position
Dribbling	Ball Control, Dribbling, Reaction
Passing	Vision, Crossing, Short Pass, Long Pass
Pace	Acceleration, Speed, Agility
Defense	Marking, Sliding Tackle, Standing Tackle, Interceptions

Table 2: Features of representing personal skills

chart charts, which express the attributes, relationships, etc. that we want to show to the user, but there are problems such as duplication.

7 DISCUSSION

For the sake of visualizing the personal profiles, we implement radar charts and line charts. Since each player has too many attributes, we need to integrate them and extract some hidden information. Therefore, we utilize six main features: physical, shooting, dribbling, passing, pace, and defense, to show each player's skills. Each feature consists of several attributes, as shown in 2, and we take the mean of those attributes to get feature values. Both radar charts and line charts show information about personal skills, but they have different intuitions. Radar charts show comprehensive information. The larger the area of the hexagon, the stronger the skill of the player. If the hexagon has sharp edges, it indicates that the player's certain skill is more prominent than the other skills. Due to several players' personal information showing in the radar chart at the same time, when we show multiple players' information on the radar chart at the same time, target users are unable to obtain accurate data due to the overlapping lines. To solve the problem, we add some interactions to the radar chart. When a user moves the mouse to the corner of the hexagon, the chosen player's hexagon will be filled with the representing color. Additionally, The exact value of the skill will be shown. However, radar charts only show the overall skill of the player, and we cannot directly see the specific value of each ability from the charts. Therefore, we add line charts, which can be used to compare the individual indicators. The value range is from 0 to 100, and the larger the value, the stronger the skill of the player.

To rank the highly-rated players according to their position and attribute, we generate a bar chart. As shown in Fig.5, we provide several buttons like positions, attributes, and detail skills. For example, if a user chooses 'Any' position and 'Rating' attribute, the top 20 players will be shown in the bar chart in decreasing order. The bar chart can display relative numbers of players by clicking 'More' or 'Less' buttons. Since one bar chart shows the same position and attribute, we didn't add any hue to the chart.

The scatter chart is used for revealing two attributes. It shows the relationship between two variables and is the best method to show a hidden pattern. The range of data flow in each scatter point is visualized by hue.

8 CONCLUSION

Association football is the most popular sport in the world, and football fans want easily and quickly to get information about players who they are interested in. Therefore, our target users are general football fans, who would like to know their favorite players' profiles and some rankings according to some attributes. In this project, we have two main ideas: first, extract latent information, which will be helpful for analyzing similarity among football superstars; second, visualize the data for three different target users. To tackle this issue, we provide several figures to meet users' satisfaction. We utilize FIFA19 dataset to show the information about football

players profiles, including personal information and personal skills by drawing hexagon charts and line charts. To rank highly rated top 20 players according to some attributes, we use bar charts and add some interactions to load more or less info. Scatter plots are used to reveal the correlation between two hidden attributes, e.g. height and jumping shown in 5. As a result, target users can easily dig out the significant information from the visualization charts.

REFERENCES

[1] https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html.
[2] T. Munzner. *Visualization Analysis and Design*. CRC Press, 2014.