COM-480 Data Visualization: Milestone 2

Marie Biolková, Marija Lazaroska, Ljupche Milosheski

May 7, 2021

Introduction

As outlined in Milestone 1, the goal of this project is to create an interactive tool showing the best tennis players in the history of Open Era (1968-) and their rivalries. To do so, we will visualize data from Grand Slam finals.

Features

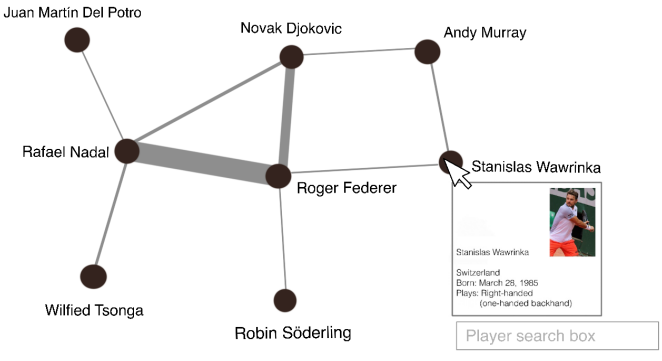
Main features

The core of the visualization will comprise of 3 elements.

- A Sankey diagram of winners and runner-ups, which will allow the user to see the most successful players in Grand Slams, since the width of the connections will reflect the number of wins (or runner ups).
- A network graph of players who played against each other in a final. Again, the edge width will be proportional to the number of occurrences. Consequently, the user will be able to see the biggest rivals, as well as identify uncommon finalists.
- Player profiles, which will show some basic information about the player. This will be displayed upon hovering over a specific player in either of the aforementioned diagrams. We would also like to link the views in these 3 components.

(a) Timeline controlled by a range slider.

(b) Sankey diagram for winners and runner-ups.



(c) Finalist network graph.

Figure 1: Sketches of main features. (Note that the data in this diagram are fictional!)

The data inputs for the above graphics (in Figure 1) will be controlled with a slider on a timeline, so that one can easily decide to display more or less information. The user will be able to specify both the lower and upper bound on the year. In addition, we want to implement a search tool to help users find players in the network. Finally, switching to data about women's finals will be possible by clicking on a button.

Sketches of the intended design can be found in Figure 1.

Extra features

There are many ways the main features can be extended to improve user experience. We will be aiming to make the links between nodes in the network graph clickable, which should then display a summary statistics of matches between the given pair of players. Unfortunately, for such a feature we were only able to scrape data about men's finals – this, together with the fact that it is not essential to what we are trying to show with this visualization, is why we will only consider it in case we have enough time left. In addition, to make the network easier to navigate, we would like to implement fish-eye zoom on the nodes.

Prototype

Here is a working prototype of the website which includes the 3 main components of our visualization mentioned before.

- We implemented a working Sankey diagram where all the winners are placed on the left and the runner-ups on the right. In the middle we have the four Grand Slam tournaments which inform the user for which tournament the players were finalists/runner-ups. Currently the tournament nodes don't include the tournament logos. In addition, we implemented the hovering event for the links which shows the number of losses/wins of the players.
- The network graph includes all the players, which are represented by the nodes, and the links between the finalists that played together. In the current prototype the links between the nodes don't have the corresponding width to the number of times the players played against each other which is something that we need to work on. Furthermore we implemented the hovering ability over nodes that creates a box where the information (like a picture and the country) of the hovered player will be displayed. Now we only have the name of the player but in the next steps we are going to include the pictures and the corresponding data.
- The timeline slider we implemented works and based on the selected range of years the Sankey diagram and the network graph change their data. We also already have the two buttons that enable the user to switch between male and female data but have yet to implement a working search bar.

Finally, the current structure of the prototype is not the final one. We need to work more on the placing of the implemented elements and fixing their size, since when choosing a wider range of years there is an overlapping in the Sankey Diagram.

Tools

As a starting point to implement these things we will use many existing D3 classes such as d3.force for the network, d3.sankey for the Sankey diagram, d3.brush for the timeline. We will have to adapt them to our application and implement some modifications that we described in the previous section.
Lecture 5 will be helpful about choosing the right ways the users would interact with the data, such as the brush for the timeline slider. Lecture 6 will help us choose the layout of our visualizations and use user-friendly color schemes taking into consideration how the human visual system works. Lecture 7 is about the design and covers the different types of the search functionally, which we will implement. Lecture 10 will be important for our network as we will learn about force-directed layout graphs.