Data Visualization Milestone II

Costanza VOLPINI, Eleonora ROCCHI, Wei JIANG
EPFL, Switzerland, April, 2020

The main goal of this project is to help students to select courses at EPFL. We will provide general information to give students an overview of courses and give customized suggestions by combining it with their personal data. So our visualization contains two parts: general information and personal choices. Since there are huge amount of courses, we allow users to select the courses they want to explore and delete what they are not interested in each part.

I. General Information

We will show the trend of courses and compare their popularity through the changes in the number of student enrollments over the years. Apart from that, we will show some additional statistics of the courses to provide more information, such as the professor, the distribution of student sections. We choose bubble diagram as our visualization to show the changes of student enrollments in each course and also compare popularity among different course in every year, as shown in Fig. 1. By clicking on a course's bubble, a chart with statistics of that course will be shown to the users.

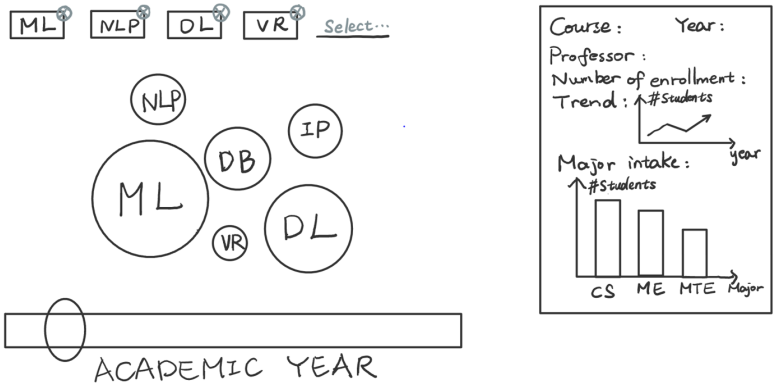

Fig. 1: Bubble diagram

II. Personal Network

We will create the course network for users to give suggestions on course selection. If the users insert their names, the previous courses they took will appear so they do not need to input all the course names. Also, we allow them to delete the course they do not like and add the courses they will take in the future. The network will appear automatically with the courses they choose and the ones we recommend. Our suggestion is based on the most popular choices among students. For example, if many students taking course A and B also take course C, we assume C is strongly connected to A and B. By clicking the course node, the users can get the relationship graph to show its top related courses. We will use Force-directed graph as to show the network and Chord to show the relationship when the user clicks on a course. An example is showed in Fig.2. The recommended course nodes build a strong network together with previous courses.

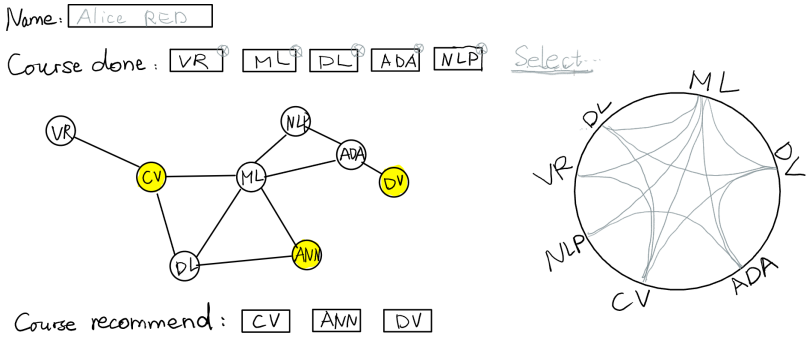

Fig. 2: Network diagram and chord chart.

III. Tool & Implementation

We have decided to use *Node.JS* an open-source and cross-platform JavaScript runtime environment. To make everything easier and simplify our development we have decided to use also the framework *ExpressJS*.

Fig. 3: Schema of the webapp project

All the data are hosted on a *MongoDB* cluster. As front-end libraries we are using *bootstrap* for styling and *d3.js* for the visualization. Last, the website is hosted on *heroku*. The web-application is located on `webapp` folder. The structure of the code inside the folder looks like Fig. 3. Where app and models are folder containing respectively the first the folders css, images, js and the second all the models for the database. The main idea was to execute a query to retrieve the data for our visualization by a query to *mongoDB* and then pass the data post-processed to *d3.js*.
The lectures we found most useful were the ones about *d3.js* and *perception colors*, even if the latter has not been applied in our prototype. Then, we look forward for the lecture regarding graphs, we think it will help us in the development of the personal course network.

IV. Extra Ideas

1) **Color of bubble in the general information diagram** We can improve our bubble diagram by adding meaning to the color of bubbles. For example, the color can show the similarity of courses. The similar courses have similar colors, which give users more information.
2) **Add the course's semester in the personal network** We can allow users to choose the Spring/Fall semester they want to take courses in, we think this way our recommendation will be more useful. (e.g. one last semester left, need for suggestions only for Fall or Spring).
3) **Add information on who is taking the course** For each course, we can show the proportion of PhD students and master students. Among master students we can show the proportion of their master year, i.e, whether they take the course in first year or second year. We guess the course with many PhD students and second-year master students are relatively more difficult. So, this information is helpful for students to select courses.
4) **Comparison to other students** We allow users to insert the name of students whom they know has similar specialization to them. Then the comparison graph will show to users to give them a hint on course selection.
5) **Course pass rate** We can roughly estimate the course pass rate through calculating the proportion of students who retake the course to give user some hint of difficulty of the courses.
6) **Change of professors** We can show the frequency of change of professor in one course and show the effect of change through the increase or decrease of the number of enrolled students.

V. Prototype Review

Link of our prototype: https://xiaozuepfl.herokuapp.com/.

Our "ISN'T ACADEMIA" website structure borrows the style of "IS ACADEMIA" which is official course selection website at EPFL. Our prototype will enable students to better visualize some course statistics.
There are 3 tabs in our website:

- **Home**
  It will contain some description of our project and dataset. Also, we will briefly introduce the functionalities of our website so users can follow the instruction and use it quickly.
- **Explore**
  Users can get general information of course by clicking the "Explore" tab and an random example will show to them. In future, they will be able change the courses to what they want to explore by inserting and deleting the course names. Also, they will be free to move the sliding year bar to know the number of enrollments of courses in a certain year and the changes through observing the expanding or shrinking of the course bubbles. Moreover, when they will click the bubble and some statistics of the chosen bubble course will show.
- **Personal**
  It will contain the 'Personal Network' graphs. Now, the users can insert their names (surname + name format) and then the courses they did before will show to them (dark blue) with some suggestions for new courses (light blue). In future, they will be free to delete or insert other courses. Also, by clicking on a course node, the relationship graph(Chord diagram) will show to give them an idea of relation degree of each courses. We are considering to reduce the degree of the nodes in our graph and to put in the chord diagram more 'related courses' than in the graph as the visualization of a graph with more edges is less enjoyable and more confusing. In the prototype, the chord plot is not connected with the nodes, we just show a sample of how the diagram will look like.