Report for Project

Field R-Layer

Zhou Mengdie

5130309134

SJTU

School of Information Engineering

June 22, 2016

Abstract

This report will introduce my project that has been finished within this semester. The name of my project is "Field R-Layer". R means round. As soon as you see the example later, you would definitely understand it. And Field R-Layer is a hierarchical classification of all academic fields, which is quite useful for both beginners and scholars.

Contents

Page

1 Motivation **3**

1.1 Review 3

1.2 Background 4

2 Preparation **6**

3 Performance **7**

4 Function **9**

4.1 For us 9

4.2 For professors 9

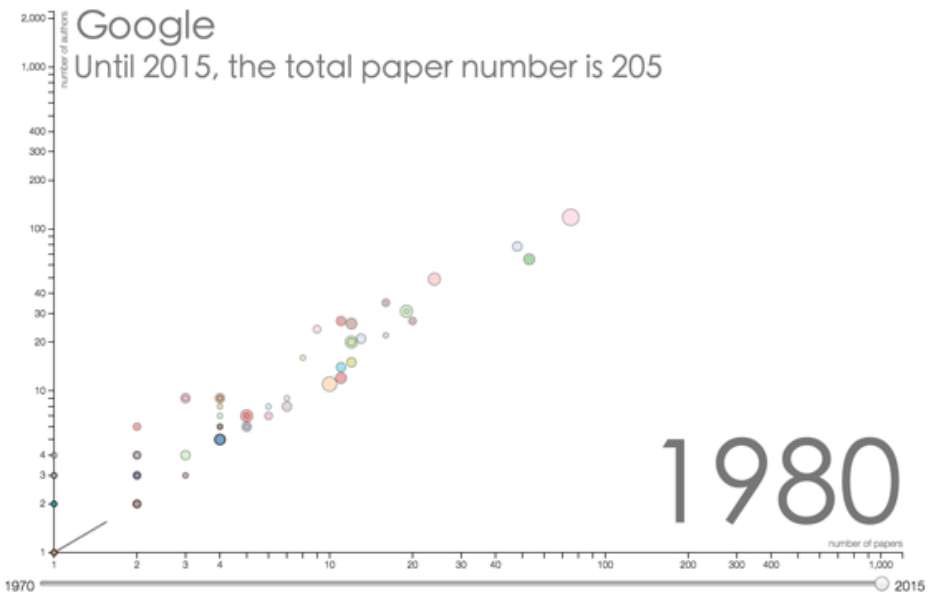
5 Future Work **10**

References **10**

1 Motivation

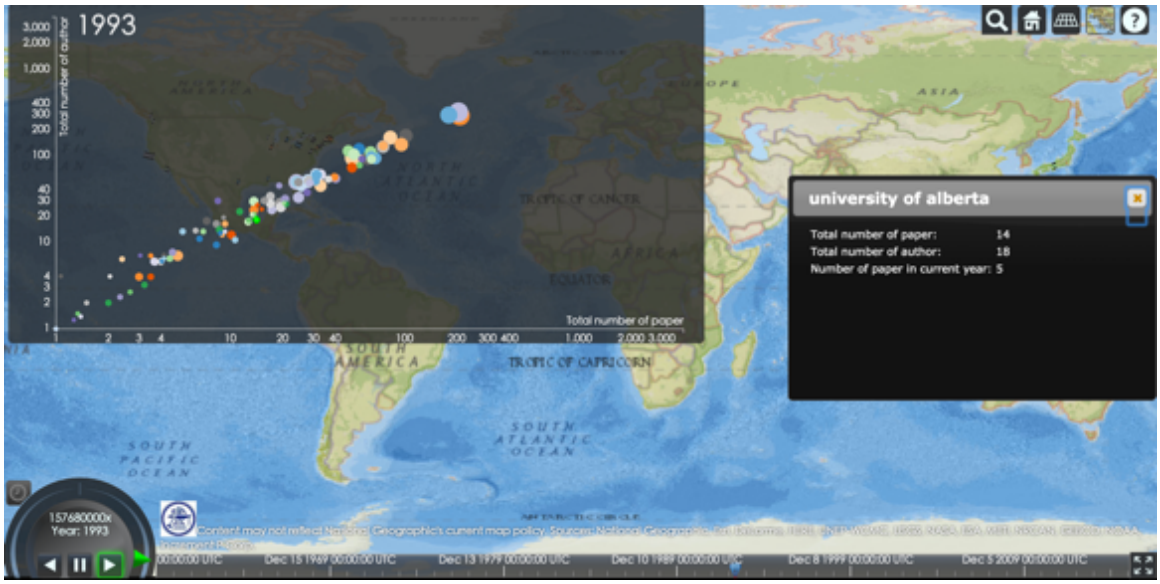
1.1 Review

Field R-Layer is the third visualization of our group. We have already done schoolDrag and schoolMap. And I would show a part of our former work.



SchoolDrag

SchoolDrag shows the trajectory of one institution or one school in relation to the number of papers and authors.Of course, we can only compare them in one field.



SchoolMap

SchoolMap take the geological factor into our consideration, which make the pattern more convincing.

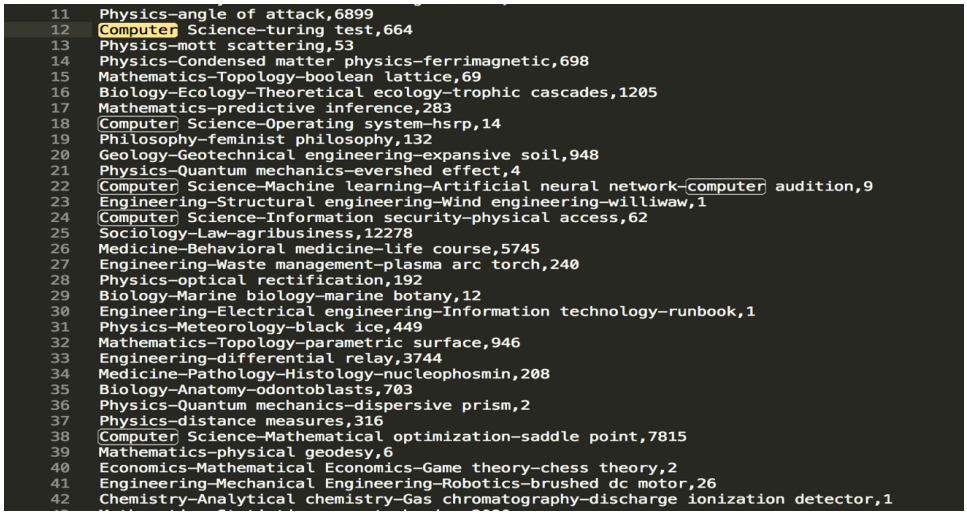
1.2 Background

However, no matter schoolDrag or schoolMap, they only focus on one field, what about other fields? How people choose different fields to match different needs. By then, we realized we lack a bigger picture which functions as a directory. Here comes the idea of round layer that differs with any other traditional directories.

2 Preparation

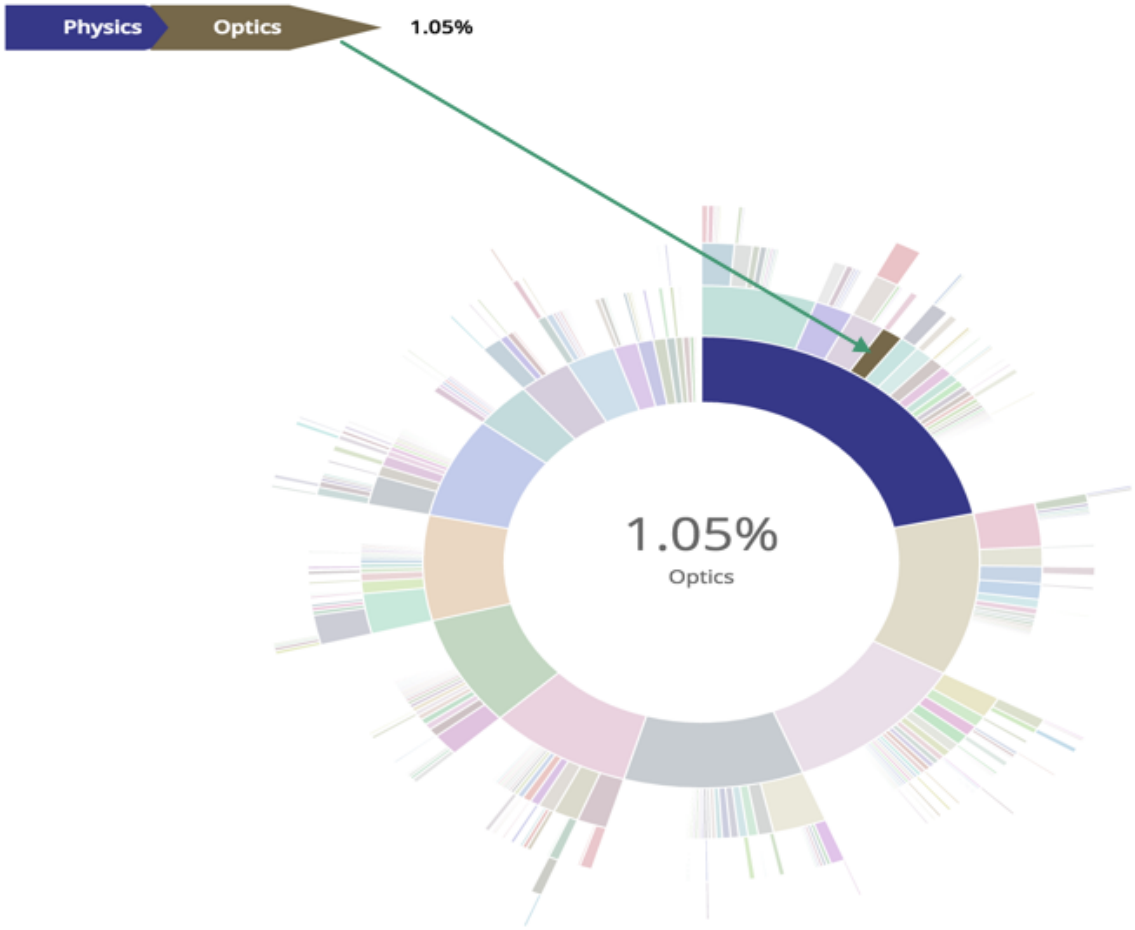
According to the picture I just showed, we know that one section has its root and children, I use several tools to get that relationship among every line of fields.

MySQL is a tool used for crawling data we need from the database. Then we use python to transform the data original format into what we need which I will show next. Finally, we us javascript to visualize our data as a whole, and D3 library is quite necessary.



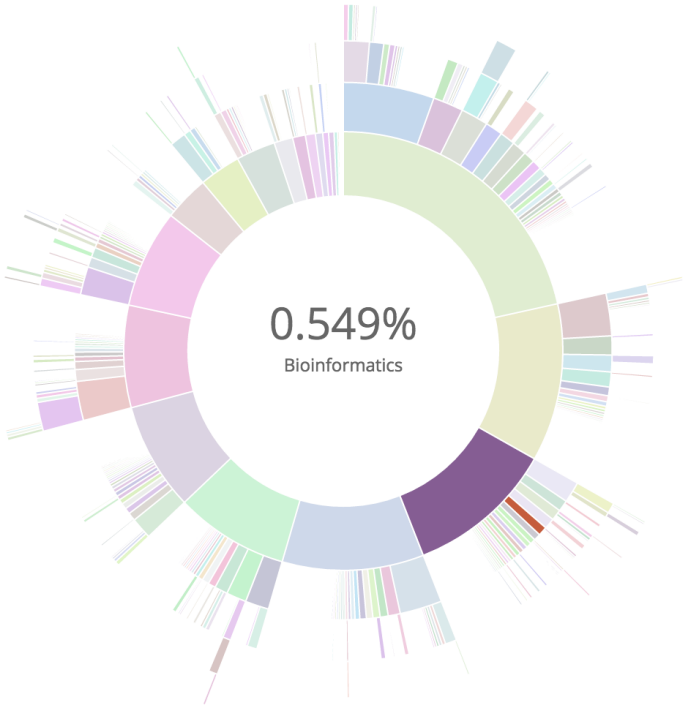
Here is the format of our data. It is in csv form. The numbers listed is the amount of papers in the leaf field. In order to obtain the percentage of every field, we have to traversal all the data, and add the numbers together in consistent parent-field. Same leaf would never occur again. To have a better visualization, I rule out the fields that the percentage is smaller than 0.1 percent. First term is the root field.

3 Performance



The picture listed is a tiny example of my project. I put the arrow on purpose so that you can identify which section has been chosen. The chosen section would show deeper color and all the unchosen sections would show lighter color. We only generate random color to distinguish the fields and show the percentage and name of the section that our mouse move over. Moreover, all the sections was in a descending order so you can see which field account for most proportion, physics obviously.

Here I will give another example of our most related field, computer science.



From this picture, we know that computer science is the third largest field among all academic fields.

4 Function

I believe most of you must have seen some other hierarchical demos such as the cluster graph, they use the size of a circle to represent the influence, which is more dynamic. But our work is much clearer especially for the beginners and operated easily.

4.1 For us

For us, we could basically have a bigger view(overview) of all academic fields and if you are more interested in one field, click on the section, then the webpage may jump to schoolDrag or SchoolMap to show other features in this field. That function as a reference for those beginners and link all the three works we've done. But it request us to finish the real-time computing first.

4.2 For professors

For professors, they could compare the proportion of their related research field with other competitive fields.That may encourage them to progress more faster in a competitive environment.

Job hopping (joke)

5 Future Work

In the future, we could take advantage of the depth of color to identify the potential of a field. This work contains a lot of algorithms that's worth discussing.

Our database sometimes use keywords rather than a field name. It could be confusing to some users and it lower the accuracy of our project.

And there are more shortcomings yet to find.