

EE101 Final Project Report

Yifu Chen, Jialong Guo, Ziliang Guo, Aofan Jiang

June 23, 2019

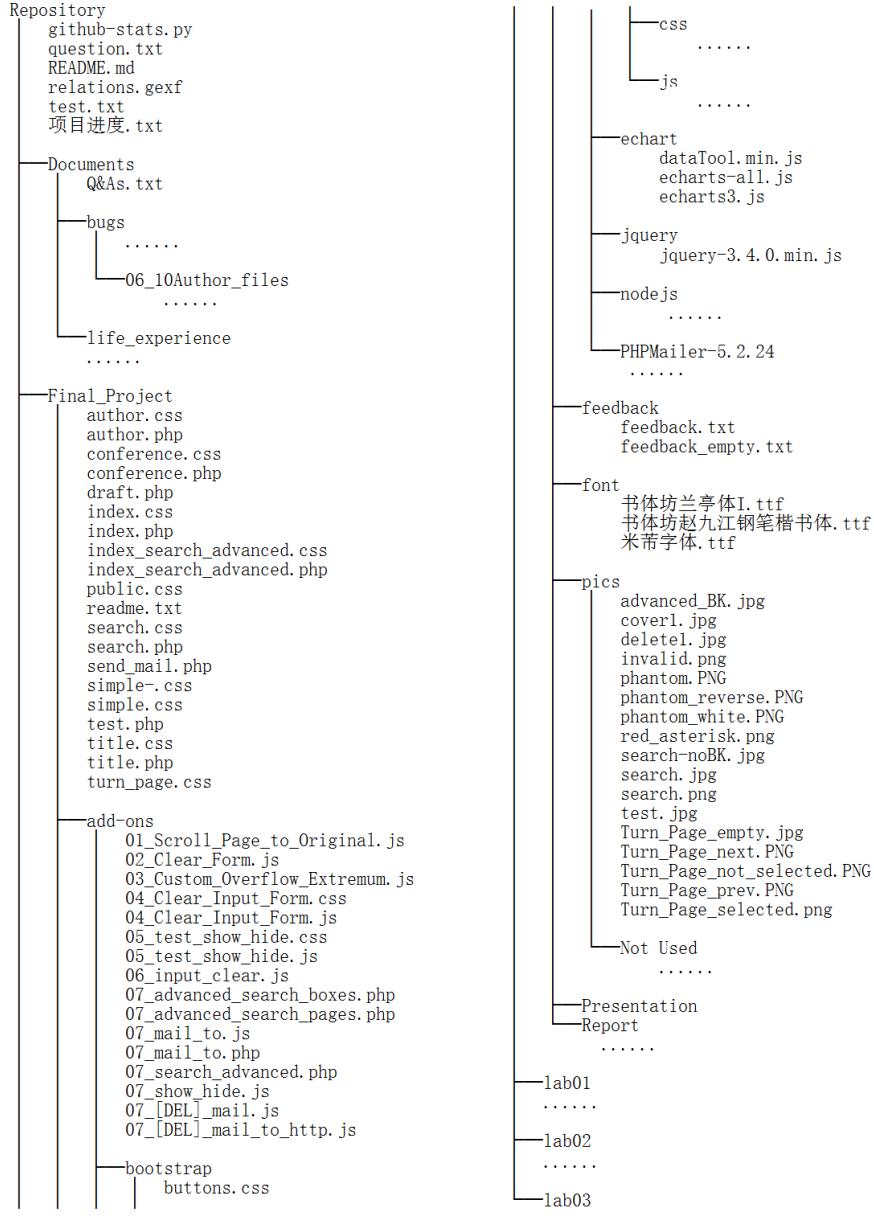
content

1 Overview	1
1.1 Project File Tree	1
1.2 Github Repository	1
1.3 Contributions	2
2 Front-end	3
3 Graphs	3
3.1 Overview	3
3.1.1 Basic Idea	3
3.1.2 Process of development	3
3.2 Index.php	4
3.2.1 Logo	4
3.2.2 Greeting Words	4
3.2.3 Search box & buttons	4
3.2.4 Copyright & Related Websites	7
3.3 Search.php	8
3.3.1 Navigation Bar	8
3.3.2 Body Part & Text	8
3.3.3 Paper Turning & "Jump to" Button	9
3.4 Conference.php	9
3.5 Author.php	10
3.6 Title.php	11
3.7 Acknowledgment	12
4 Graphs	14
4.1 Overview	14
4.2 Design	14
4.3 Searching Data	15
4.3.1 From Mysql to Solr	15
4.3.2 Searching in Solr	16
4.4 Formatting Data	16
4.5 Drawing Graph	17
4.5.1 Collunmn Graphs	17
4.5.2 Pie Graphs	18
4.5.3 Force Directed Graphs	19
4.6 Fruits' Display	20
4.7 Codes	22

5	Introduction	24
6	Keyword Highlighting	25
7	Page Turning	26
7.1	Brief Introduction	26
7.2	Details	27
7.2.1	V1.X	27
7.2.2	V2.0	27
7.2.3	Advanced	29
8	Search - Simple and Multi-field	29
9	Webpage Features	31
9.1	Refresh	31
9.2	Click Hyperlinks	32
9.3	Page Turning Anchors	32
9.4	Loading Speed Improvements	32
10	Paper Recommendation	33
11	Advanced Search	35
11.1	Overview	35
11.2	Search Bar	35
11.3	Submit	36
11.4	Page Turning and Jump-to	37
12	Feedback	39
13	Hyperlinks	41
13.1	Hyperlink of each title	41
13.2	Hyperlink of each conference	42
13.3	others	42
14	Code optimization	43
14.1	Solr	43
14.2	Mysql	43
15	Contact and Open source	44

1 Overview

1.1 Project File Tree



1.2 Github Repository

Visit: <https://github.com/marridG/EE101-Final-Project>

or Search: SJTU 2019 S1 EE101 Final Project

1.3 Contributions

The whole project is developed during May 12th and June 10th. The detailed process is documented by Github as follows. It should be noted that the extra merge commits are excluded.



The next picture shows the contribution of each members in our group. It must be pointed out that the code commits is just a reflection of contribution but not all. “chyf0726” represents Yifu Chen. “marridG” represents Ziliang Guo. “Sjtugjl” represents Jialong Guo. “jjjaaafff” represents Aofan Jiang. The picture shows as follows.



2 Front-end

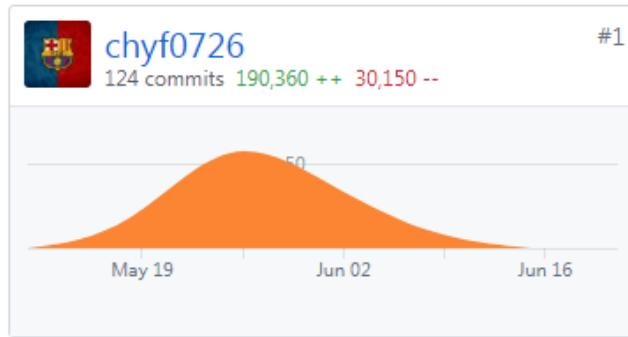
[Start] — By Yifu Chen 518030910270 —

2.1 Overview

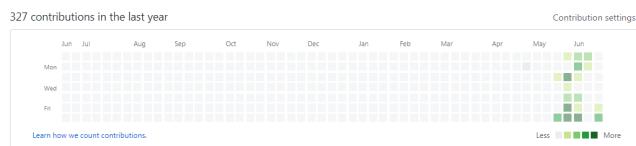
2.1.1 Basic Idea

I was in charge of the front-end part. I mainly used CSS and BOOTSTRAP to beautify our websitek. In my opinion, I hope my websites be plain and straightforward, so I did not decorate our websites deliberately, and this my idea of designing the layout fo our websites.

2.1.2 Process of development



We took the advantage of *Github* to promote our cooperation. During the period, I wrote down about 200,000 lines of codes which was mostly written independently, and I was the number 1 contributor in our team.



I started my job on 26th May and ended on 8th June and I will introduce my job in the following part.

2.2 Index.php



Our index page is shown above. First of all, I would like to introduce the process of my designing our home page. At first, there were three search boxes in the page. The search box "Author", "Title" and "conference", and the layout of the index.php was settled. Then, we decided to use multi-searching. Therefore I cut down the number of search boxes into one. Finally, I polished the index.php and the page became what you can see now.

The elements the index page consists of was shown in the graph above, and I am going to introduce every part respectively.

By the way, the favicon of our websites was *Raffaello's The School of Athens*.

2.2.1 Logo

The name of our search engine came from *The Phantom of the opera*, one of my favorite films, and I hope that the speed of our searching engine can be as fast as a phantom. It took me so much time to find an appropriate font to display our logo. Finally I found out a font "书体坊兰亭体". However, the words could not be shown in terms of vector graph, which meant that the edge of the words were not smooth, and that is a defect of our logo.

2.2.2 Greeting Words

To display the greeting words, I also spent plenty of time to find an appropriate font. Finally, I found a font named "ChannelSlanted2" to present the greeting words.

2.2.3 Search box & buttons

The search box and the buttons are the key part of the page, and I beautified the buttons, and the result was shown in the following pictures.

<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching
<input type="button" value="SEARCH"/>	<input type="button" value="RESET"/>	Advanced Searching





2.2.4 Copyright & Related Websites

These two parts were mainly written by Ziliang Guo and Aofan Jiang and Hyperlinks were set on the texts "Contact Us" and "Open source".

2.3 Search.php

I am going to introduce every part of search.php respectively.

Multi Field Search: jiawei han

Title	Authors	Conference
mining heterogeneous information networks	jiawei han;	SIGKDD
mining frequent patterns by pattern growth methodology and implications	jiawei han; jian pei	SIGKDD
mining heterogeneous information networks a structural analysis approach	yizhou sun; jiawei han;	SIGKDD
disograph mining closed frequent graph patterns	xileng yan; jiawei han	SIGKDD
dynamic generation and refinement of concept hierarchies for knowledge discovery in databases	jiawei han; yongjian tu;	SIGKDD
chinese-japanese cross language information retrieval a han character based approach	yujii matsumoto; midmaru hasan	ACL
resource and knowledge discovery in global information systems a preliminary design and experiment	osmar r. zaher; jiawei han;	SIGKDD
on trivial solution and scale transfer problems in graph regularized nmf	chris ding; jiawei han; quanquan gu	ICAI
advances of the oblearn system for knowledge discovery in large databases	simon tang; jiawei han; yongjian tu;	ICAI
classifying large data sets using svms with hierarchical clusters	jiong yang; jiawei han; huiqiao yu	SIGKDD
an efficient multi-relational naive bayesian classifier based on semantic relationship graph	xiaoxin yin; jiawei han; hongyan liu;	SIGKDD
collection topic modeling for heterogeneous networks	jiawei han; bo zhao; hongbo deng	SIGIR
tensor space model for document analysis	jiawei han; deng cao; xiaolei he;	SIGIR
clustering moving objects	yitao li; jiong yang; jiawei han	SIGKDD
metarule guided mining of multi dimensional association rules using data cubes	micheline kamber; jiawei han; jennifer chang;	SIGKDD
robust tensor decomposition with gross corruption	huan qui; jiawei han; quanquan gu	NIPS
trust analysis with clustering	manish Gupta; jiawei han; yizhou sun;	WWW
ranking event periodicity from incomplete observations	jingjing wang; zhenniu li; jiawei han	SIGKDD
ranking based clustering of heterogeneous information networks with star network schema	jiawei han; yizhou sun; yintao yu;	SIGKDD
ranking based classification of heterogeneous information networks	mariia danilevsky; jiawei han; ming ji	SIGKDD
building enriched web page representations using link paths	tm weninger; jiawei han; chenxiang zhai;	WWW
parallel mining of closed sequential patterns	jiawei han; david padua; shengnan cong	SIGKDD
spectral regression for efficient regularized subspace learning	jiawei han; deng cao; xiaolei he;	ICCV
sparse projections over graph	jiawei han; deng cao; xiaolei he	AAAI
closed set mining for the best strategies for mining frequent closed itemsets	janyong wang; jiawei han; jian pei	SIGKDD

localhost/EE-101-Final Project/Final Project/title.php?title=mining%20frequent%20pat...&with=clustering

Welcome To Phantom Academia Searching

Found 401 results. Each page: 25 items. Altogether: 17 pages.
Text

Phantom

Page turning: 1 2 3 4 5 6 7 8 9 10 Next

Jump to: Go

2.3.1 Navigation Bar

I used css and bootstrap to construct the navigation bar. I thought there was no need to build a that complicated navigation bar, so I just put our logo in the top left corner of the screen and set it with a hyperlink to the homepage, and in the top right corner if the screen was a search box to conduct multi-search.

2.3.2 Body Part & Text

I selected the font "Regencie" to beaufy the table and "书体坊赵九江钢笔楷书" to beatufy the texts.

2.3.3 Paper Turning & "Jump to" Button

Our idea to design the pattern of the Paper Turning is to imitate *Google's* pattern.



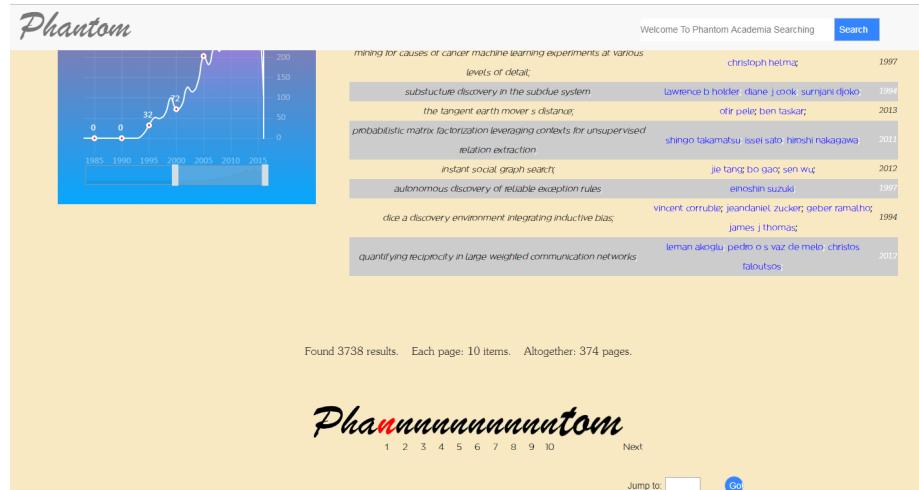
Therefore, I photoshopped some pictures and realized this idea.

I beautified the "Jump to" box, and the button "Go!" was modified by Aofan Jiang.

2.4 Conference.php

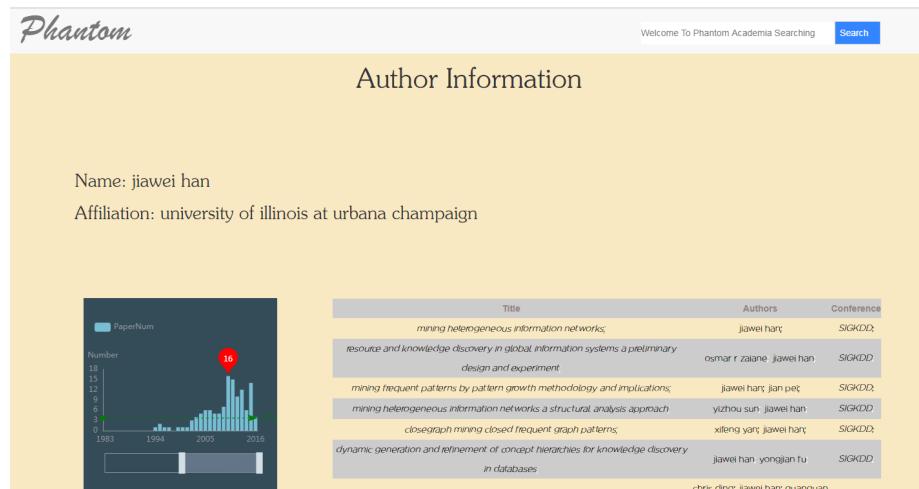
The screenshot shows a search interface for the SIGKDD conference. At the top, there is a search bar with the placeholder "Conference Name: SIGKDD". Below the search bar is a chart titled "Publication - Year" showing the number of publications over time from 1985 to 2013. The chart has a purple-to-blue gradient background. To the right of the chart is a table listing conference papers. The table has columns for "Title", "Authors", and "Year". The first few rows of the table are:

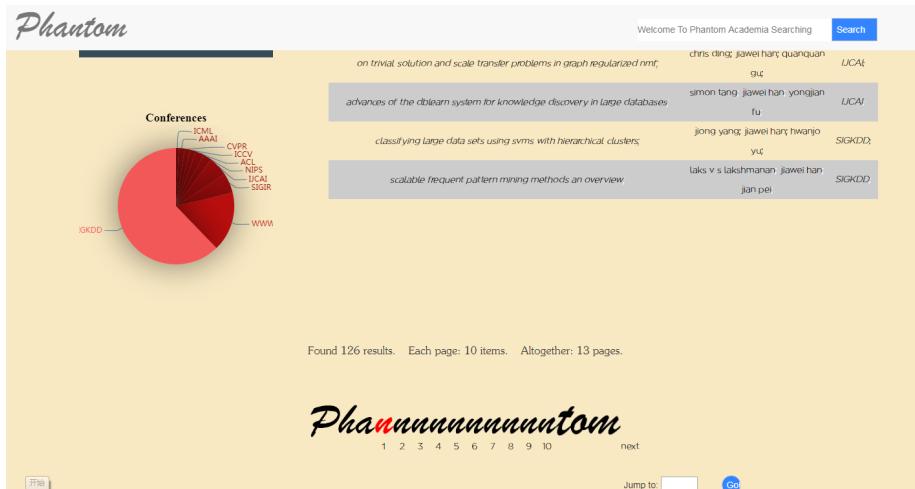
Title	Authors	Year
semi supervised parameter free divisive hierarchical clustering of categorical data using the bhattacharyya mean for the filtering and clustering of positive definite matrices	andré mayrger ernest moniq shengru wang tengfei xiong	2011
mining for causes of cancer machine learning experiments at various levels of detail;	christoph helming	1997
substructure discovery in the subdue system the tangent earth mover's distance;	lawrence b holder diane j cook surjanari djoko otti peter ben tsakas;	1994 2013
probabilistic matrix factorization leveraging contexts for unsupervised relation extraction	shingo takamatsu issei sato hiroshi nakagawa	2011
instant social graph search	jie tang bo gao sen wu	2012
autonomous discovery of reliable exception rules	enoshin suzuki	1997



The beautification of Conference.php was similar to the search.php. Therefore I will not go into details here.

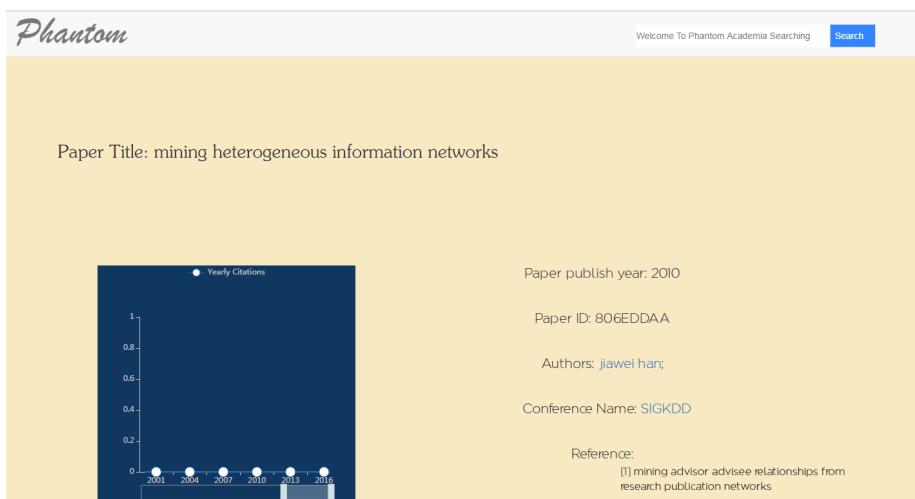
2.5 Author.php

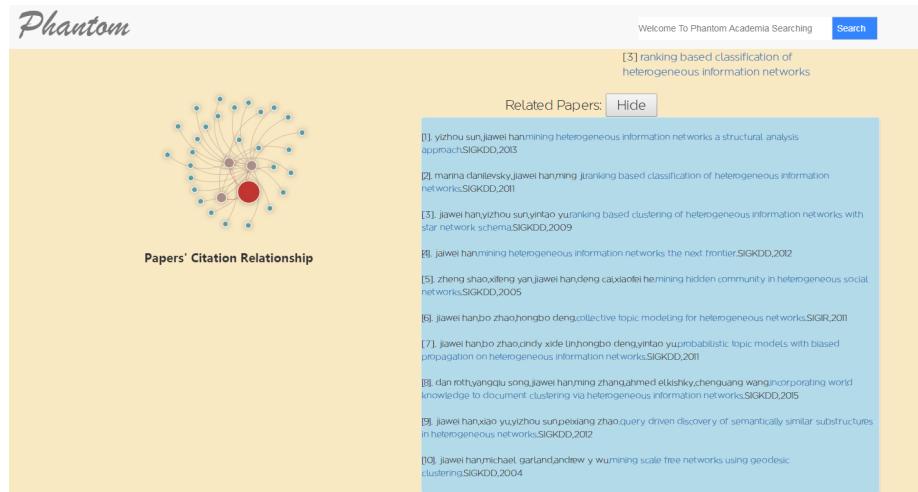




The beautification of the page author was similar to the search.php. Therefore I will not go into details here as well.

2.6 Title.php





I set the texts of the website in a conspicuous position, and the all the texts in the page were shown in selected fonts. I beautified the "Show" button as well.

2.7 Acknowledgment

I would like to talk about what I learnt after taking the course *Introduction to Engineering for Electronic Information (B)*. Well, I have to say that this course is a little challenging. The 3 labs and the final projects all covered the fields that I was not familiar with or even never heard about. All these tasks needed me to solve tons of problems and many of them made me mad. However, I got through all these problems with my own effort and the help from my roommates, my classmates and TAs. So here I want to say "Thank You" to all that helped me. I especially want to express my gratitude to TA *Hui Xu*. It was his help and patience that benefited me a lot. I remembered that I met a problem in lab3, and I spent about 4 days solving the problem, but it was still unsolved. Then I turned to TA *Xun*, and we solved the problem together and finally worked it out.

Then I would like to talk about my work in the final project. As my roommate took the compulsory parts of the final projects, I "had to" do the beautification. I do not think beautification is a job of difficulty, but it required patience. I usually spent plenty of time searching for a awesome fonts and racked my brains to think how can do the webpage better, but when writing this report, I really found it difficult to present my job, for what I have done are only about the layout of the website and some trivial details. I cannot write something like "I used xxx to quicken the search" or "I modified my algorithm with this bright idea". Beautification work is trivial, time-consuming and requires patience but I think I really learn a lot form the work.

Finally,I want to say sorry for my absence of the last course.I do not want to explain anything.The only thing I want to clarify is that I never have any disrespect to Professor Wang and all TAs, and I feel regretful for my absence and apologize sincerely.

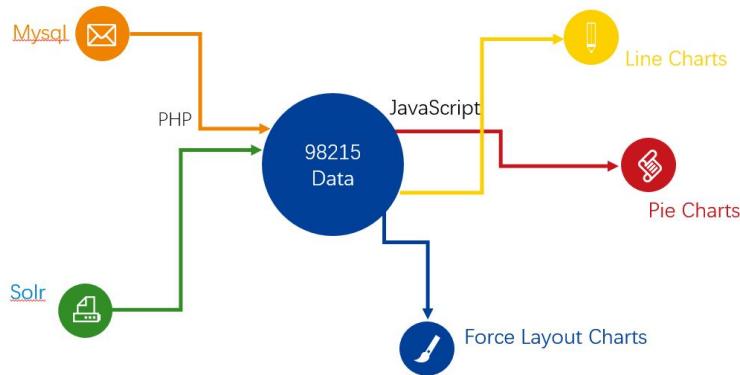
Thanks again to all those who have helped me!

[End] — By Jialong Guo 518030910273 —

3 Graphs

[Start] — By Jialong Guo 518030910272 —

3.1 Overview



When it comes to our graph part, it can be chiefly divided into three parts:

The first of which is to get all the data we need in order to draw a graph.

The second part of which is to format all the data into the form that the corresponding graph is needed.

The last part of which is just to use specific javascript library to create graphs and make them beautiful.

3.2 Design

In author page's left part, I add a column graph, which displays the number of publication the author has yearly and shows the trend of his/her academic activity to give users a clear impression about the author's research fruit and the frequency of his/her delivery.

In author page's left part, I also add a pie graph, which showcases how many papers the author publishes in a conference which brings a clear view on the author is mostly connected with what conference.

In paper page's upper part, there's a line-column graph showing its yearly

citations, from which we can gain a insight into the popularity of its research field.

In paper page's lower part, I use a force directed graph to display the relations between similar papers. It gives a convience way to find related papers and messages.

In conference page, a line graph of its yearly amount of papers may reveal its academic influence.

3.3 Searching Data

In this subsection I mainly discuss the first part of graph drawing, get data from database. Since diverse graphs need diverse data, here I just demonstrate a specific example of getting data, which can mostly stand for my means of searching data.

3.3.1 From Mysql to Solr

Taking time cost into consideration, we may need to import data into solr from mysql in advance, for searching from solr will spend more time than from mysql. Thereby, it is required that we write the referenceID of reference papers into solr's schema, which allows us to search the reference papers of a paper just from solr when needed.

The final data we put in solr is formed in this way:

```
with codecs.open(FP_out, 'w', 'utf-8-sig') as f:
    data = {"PaperID": result[0][0],
            "Title": result[0][1],
            "Authors'ID": [result[0][2]],
            "Authors'Name": [result[0][3]],
            "ConferenceID": result[0][4],
            "ConferenceName": result[0][5],
            "Year": result[0][6],
            "ReferenceID": [result[0][7]]}
    #
    # "AffiliationID": [result[0][8]]}

    for i in range(1, len(result)):
        out_print = False
        if result[i][0] == data["PaperID"]:
            if (result[i][2] not in data["Authors'ID"]):
                data["Authors'ID"].append(result[i][2])
            if (result[i][3] not in data["Authors'Name"]):
                data["Authors'Name"].append(result[i][3])
            if (result[i][7] not in data["ReferenceID"]):
                data["ReferenceID"].append(result[i][7])
```

You can refer to the codes attached for detailed information.

3.3.2 Searching in Solr

The first step is getting value from user's input, then create the url link to search in solr. After this step, we can get a .json file with the result.

3.4 Formatting Data

In this section, we mainly discuss how to change the search result into suitable formation of drawing a graph. Since I searche data using php sentence, this time I have to turn php to javascript because the graph uses javascript.

Force directed graph:

Force Graph is a graph whose format is rather difficult to handle, which contains two parts. The first one is called node categories, for which we need to name every node and divide them:

```
echo "<script>
graph.nodes.push({category: 'Level One Citations', name: '$p', value: 45 , label: '$title'});
</script>"
```

The second part is line data, for which we need to give the source and target of a line as well as its value. In addition, in php foreach sentense is helpful in handle plenty of data when formatting them.

```
echo "<script>
graph.links.push({source:'0', target:'$p',value:200});
</script>";
```

Column Graph:

In column graph, an array which stores the value of yearly publication is needed. So a basic command of Javascript's array is required in order to draw a column graph. Also, the key that turn php sentence to js is echo structure, which is shown in the following picture:

```
echo "<script> var datas=new Array() ; </script> ";
echo "<script> datas[$year-1950] = \"$value \" ; </script> ";
```

Pie Graph:

In pie graph, the case is almost similar to column. I just start another array in order to put the data into it.

Line Graph:

In terms of line graph, it is also needed that a container to put all the data, which means an array is a good choice for us to use.

3.5 Drawing Graph

In the section, the process of drawing with echarts is mainly discussed. Echarts is an opensource visualization library realized by JavaScript, compatible with most current browsers (IE8/9/10/11, Chrome, Firefox, Safari, etc.) . This is a very useful tool in drawing graphs. Compared to D3.js, echarts is much more easier to learn to edit, and saving much more time in writing codes. Therefore, I chose echarts anyway.

When drawing graphs the totle step is (1) prepare a container,(2) initiate a echart, (3) draw a graph, (4) set options. The first step is just use `div` to put a graph, and easy to use. I'll introduce different ways to draw graphs and candidates.

3.5.1 Collunmn Graphs

Column graphs can show clearly the exact number of data, I chose it to showcase the yearly citations of an author.

Basically I need to set the x-axis, y-axis and the legend. What's more, I used function in realize the formatter setting :

```
formatter: function(p) {
    return p.value > 0 ? (p.value) : '';
}
```

And datazoom is imported in order to show diverse year's of publication. Users can drag it and change the length of time:

```

    "dataZoom": [
      {
        "show": true,
        "height": 30,
        "xAxisIndex": [
          0
        ],
        "bottom": 30,
        "start": 50,
        "end": 110,
        "handleIcon": 'path://M306.1,413c0,2.2-1.8,-3.2-4.5,-3.2L300,375c-1.8,0-4.5,3.2-4.5,3.2L266.1,413z',
        "handleSize": '110%',
        "handleStyle":{
          "color": "#d3dee5",
        },
        "textStyle":{
          "color": "#fff",
          "borderColor": "#90979c"
        }
      },
      {
        "type": "inside",
        "show": true,
        "height": 15,
        "start": 1,
        "end": 35
      }],
    ...
  ]
}

```

Moreover, I set functions as well to find the maximum and average of data

```

  ...
  "markPoint": {
    data: [
      {type: 'max', name: '最大值', itemStyle:{color: 'red'}}
    ]
  },
  markLine:{
    data:[
      {type:'average',name:'平均值',itemStyle:{
        normal:{
          color: 'green'
        }
      }}
    ]
  },

```

3.5.2 Pie Graphs

Pie graphs can display the ratio of a part to a total part. Since there are static amount of conferences, I just check the number of each conference to draw the graph, and then eliminate the zeros.

```

series: [
  name: 'conference',
  type: 'pie',
  radius: '60%',
  center: ['50%', '50%'],
  data: [
    value: datas2[0] == 0 ? null: datas2[0],
    name: 'AAAI',
    // itemStyle: createRandomItemStyle()
  ],

```

3.5.3 Force Directed Graphs

Force directed graph can be divided into lines,nodes and lables.

Lines' setting can change their length based on their correlation degree, the curvature and color.

```

lineStyle: {
  normal: {
    show : true,
    color: 'source',//决定边的颜色是与起点相同还是与终点相同
    curveness: 0.3//边的曲度, 支持从 0 到 1 的值, 值越大曲度越大。
  }
},

```

Nodes' setting can change their size,color, etc.

```

itemStyle: {
  normal: {
    borderColor: '#fff',
    // radius : 100,
    borderWidth: 1,
    shadowBlur: 10,
    shadowColor: 'rgba(0, 0, 0, 0.3)'
  }
},

```

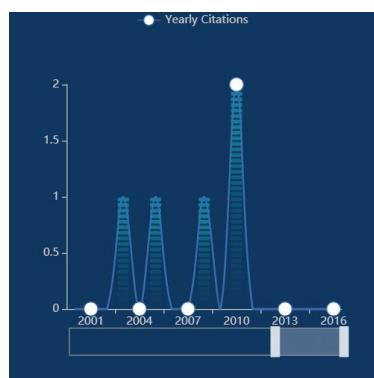
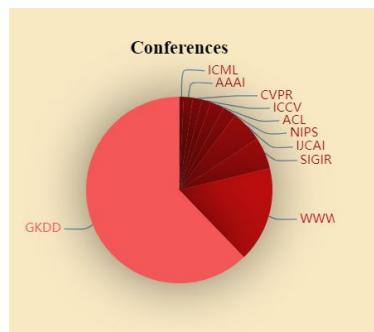
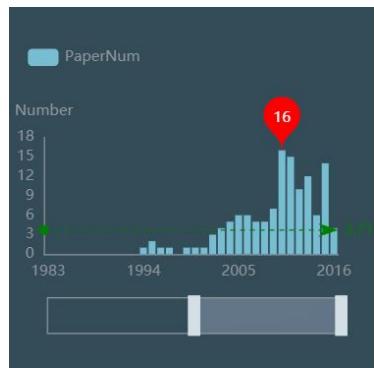
Lables can make data more easy to different.

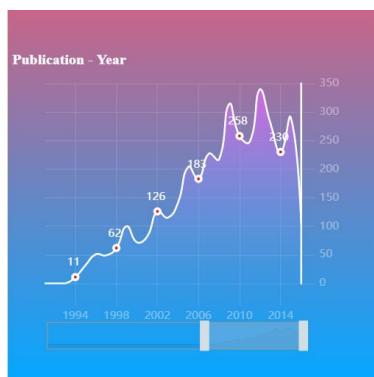
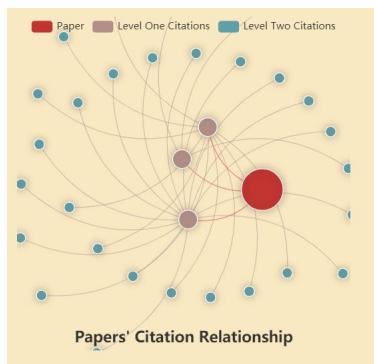
```

label: {//图形上的文本标签, 可用于说明图形的一些数据信息
  normal: {
    show : false,//显示
    position: 'right',//相对于节点标签的位置
    //回调函数, 你期望节点标签上显示什么
    formatter: function(params){
      return params.data.label;
    },
  }
},

```

3.6 Fruits' Display





3.7 Codes

```
$ch = curl_init();
$timeout = 5;

    //Search for specified year's paper number in a conference.
echo "<script> var datas=new Array() ; </script> ";
for ($year=1950; $year<=2016 ; $year++) {
    $keyword = $year;
    // $author_id =
    $query1 = urlencode(str_replace(' ', '+', $keyword));
    $query2 = urlencode(str_replace(' ', '+', $author_id));
    $url = "http://localhost:8983/solr/lab02/select?q=Year%3A%20$query1%20%26";
    curl_setopt ($ch, CURLOPT_URL, $url);
    curl_setopt ($ch, CURLOPT_RETURNTRANSFER, 1);
    curl_setopt ($ch, CURLOPT_CONNECTTIMEOUT, $timeout);

    $data = json_decode(curl_exec($ch), true);
    $value = $data['response'][ 'numFound'];
    echo "<script> datas[$year-1950] = \"\$value \" ; </script> ";
}

    //Search for specified paper's citation frequency in a conference.
$conferences = array("AAAI", "CVPR", "ACL", "IJCAI", "NIPS", "WWW", "ICCV", "ICM");
echo "<script> var datas2 = new Array(); </script> ";
for($i = 0;$i < 13;$i++) {
    $keyword = $conferences[$i];
    $query1 = urlencode(str_replace(' ', '+', $keyword));
    $query2 = urlencode(str_replace(' ', '+', $author_id));
    $url = "http://localhost:8983/solr/lab02/select?q=ConferenceName%3A%20$query1%20%26";
    curl_setopt ($ch, CURLOPT_URL, $url);
    curl_setopt ($ch, CURLOPT_RETURNTRANSFER, 1);
    curl_setopt ($ch, CURLOPT_CONNECTTIMEOUT, $timeout);

    $data = json_decode(curl_exec($ch), true);
    $value = $data['response'][ 'numFound'];

    echo "<script> datas2[$i] = \" \$value \" ; </script> ";
}
curl_close($ch);
?>
```

```

<script type="text/javascript">
// 初始化图表标签
var chart1 = echarts.init(document.getElementById('chart1'));
var chart2 = echarts.init(document.getElementById('chart2'));
var xData = function() {
    var data = [];
    for (var i = 1950; i <= 2016; i++) {
        data.push(i + "");
    }
    return data;
}();
var option1 = {
    backgroundColor: "#344b58",
    "title": {
        "text": "",
        // x: "4%",
        left: 'center',
        textStyle: {
            color: '#fff',
            fontSize: '18',
            fontFamily: 'Times New Roman'
        },
        "tooltip": {
            "trigger": "axis",
            "axisPointer": {
                "type": "shadow",
                "textStyle": {
                    color: "#fff"
                }
            },
        },
        "grid": {
            "borderWidth": 0,
            "top": 110,
            "bottom": 95,
            "textStyle": {
                color: "#fff"
            }
        },
    },
}

```

[End] — By Jialong Guo 518030910273 —

4 Introduction

[Start] --- By Ziliang Guo 518030910273 ---

I want to highlight that I greatly emphasize the user-friendliness of my work.

(1) I took the initiative that we take full advantage of Github to accelerate our project. I also create a document to take notes of the problems we met and the solutions.

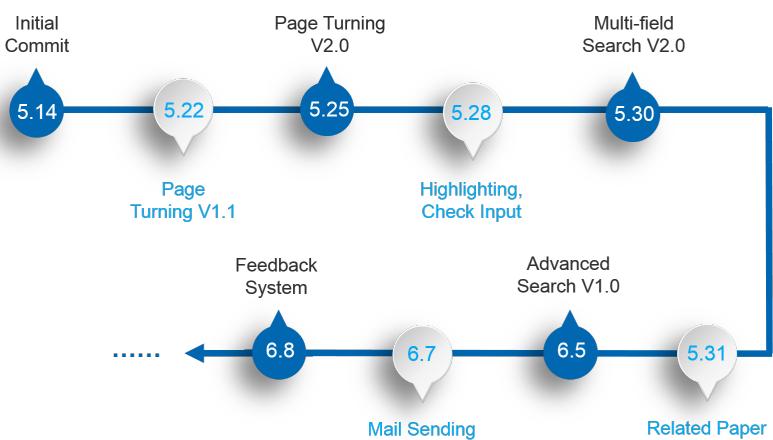
(2) Actually, I wrote the manual and uploaded my Lab 01 - 03 codes to unify the databases.

(3) I mainly focus on the back-end development.

(4) Of all my codes, I wanna highlight that approximately 85% are mainly created **independently**. For the remaining codes, modification is applied, with reference to some online blogs.

(5) Meanwhile, during my coding, I always remember to leave interfaces for my collaborators.

(6) As is vividly depicted in the timeline graph, I realized and improved different sections separately, in other words, term by term. Of course, my constant improvements are shown.



Contribution: 62 commits; 1,674,012 ++; 29,176 --.

5 Keyword Highlighting

Mainly in “search” .

I adopted the “hl” settings of Solr. It is somehow very simple. Just echo the corresponding urls will do.

However, please notice that, for multivalued fields such as Authors_Name, only the highlighted part is returned. So I made judgements in such special cases.

Also refer to: <http://www.aboutyun.com/thread-9433-1-1.html>

Effects:

Multi Field Search: jiawei han	
Title	Authors
ous information networks	jiawei han;
r growth methodology and implications	jiawei han; jian pei;
n networks a structural analysis approach	yizhou sun; jiawei han;
osed frequent graph patterns	xifeng yan; jiawei han
concept hierarchies for knowledge discovery in databases	jiawei han; yongjian fu;
ation retrieval a han character based approach	yuji matsumoto; md maruf hasan;
nal information systems a preliminary design and experiment	osmar r zaiane; jiawei han;

Codes:

```
$url = "http://localhost:8983/solr/lab02/select?indent=on&q=Title:".$query."^1+OR+Authors_Name:".$query.^"0.7+OR+ConferenceName:".$query.^"0.5&start=".($page_limit*($page-1))."&rows=".$page_limit."&wt=json&hl=on&hl.fl=Title,Authors_Name,ConferenceName&hl.simple.post=<%2Fb><%2Ffont>&hl.simple.pre=<font%20color%3D%23FF0000><b>";
```

```

// print all the Authors_Name
echo "<td>";
foreach ($paper['Authors_Name'] as $idx => $author)
{
    $author_id = $paper['Authors_ID'][$idx];
    if($author==$key_word && array_key_exists("Authors_Name", $result['highlighting'][$paper['id']]))

    {
        $author_hl=$result['highlighting'][$paper['id']]['Authors_Name'][0];
        echo "<a class=\"output_href\" id=\"$author_name\" href=\"/EE101-Final_Project/
        Final_Project/author.php?author_id=$author_id&page=1&author_affi="" target=\"_
        blank\">$author_hl</a>";
    }
    else
        echo "<a class=\"output_href\" id=\"$author_name\" href=\"/EE101-Final_Project/
        Final_Project/author.php?author_id=$author_id&page=1&author_affi="" target=\"_
        blank\">$author</a>";
    echo "<br>";
}
echo "</td>";

```

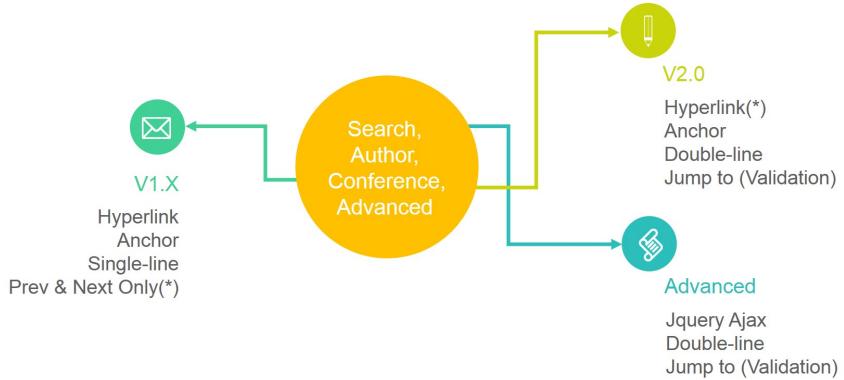
6 Page Turning

In “search” , “author” , “conference” , “advanced search” .

6.1 Brief Introduction

Actually, this part undergoes about three versions, and the V1.X version has been discarded because it is too simple.

The features of the versions are:



As for the details:

- (1) Hyperlink means that I use hyperlinks to jump to a new page.
- (2) Anchor refers to the fact that after turning pages, the page will be automatically guided to the titles of the result tables for better user experience.
- (3) Jump to stands for the jump-to function, which enables users to jump to a valid page of results.

(4) Actually, the 2.0 and higher version, I tried to imitate the page-turning function of Google. But for the methods, improved hyperlinks for the former, while Jquery and Ajax for the latter.

6.2 Details

6.2.1 V1.X

This version is just that of my Lab 03 outcomes. It is so simple that only turning to the previous and the next page is available to users.

Effects:

[Previous Page](#) [Next Page](#)

6.2.2 V2.0

In “search” , “author” , “conference” .

For this version, hyperlinks are used.

As is mentioned above, I try to imitate the page turning function of Google.

In fact, the page number of the current page and the “red-n” above it cannot be clicked and the other numbers and “black-n” s can be clicked, which will direct the user to the corresponding pages.

Meanwhile, when the next page is applicable, the “tom” , the blanked area below, the blanked area to the right and the “Next” can all be clicked. Of course, “Next” appears if and only if the next page exists. It is similar for the previous page, except for the different words (“Pha” , “Prev”).

To implement such function, just create a 2-row table and plug in the corresponding items. Also make judgements whether a page number is the current page. If not, enable cursor and hyperlinks and viceversa.

At first, I wrote a function to calculate the maximum page numbers using textbfvariable-passing-by-reference.

Actually, there are altogether four pictures here --- “Pha” , “red-n” , “black-n” , “tom” ---. They are given different file names and CSS features, so it is easy to change images.

I also added the jump-to-page function. To click the “Go” button, a valid number input is required. Otherwise, an error will be raised. The number should be a number greater or equal to 1 and less than or equal to the maximum page number.

If a floating number is inputted, say 1.1 (the same for 1.9), it will be treated as 1. Actually, I am conscious that regular expressions can be adopted here to further ensure the validation. But because of a lack of time, I failed to learn how to use it.

I found "type = number" helpful, but the default up-and-down-arrows confused me a lot. It is ugly and even disgusting!! I found a simple solution using "webkit" of CSS.

Refer to: <https://my.oschina.net/qii/blog/341439>

Effects:



Codes:

```
// Pages in the middle
for($i=$min_page;$i<=$max_page;$i++)
{
    if($i==$page)
        echo "<td><img src =\"/EE101-Final_Project/Final_Project/pics/
        Turn_Page_selected.png\" id=\"search_Turn_Page_selected\"></td>";
    else
        echo "<td><a href=\"/EE101-Final_Project/Final_Project/search.php?key_word=$
        key_word_temp&page=$i#skip_multi\"><img src =\"/EE101-Final_Project/
        Final_Project/pics/turn_Page_not_selected.png\" id=\""
        search_Turn_Page_not_selected\"></a></td>";
}
```

```

function Turn_Page_min_max_page($num_max,$page_limit,&$min_page,&$max_page,$page)
{
    if($num_max<=9*$page_limit)
    {
        $min_page=1;
        if($num_max%$page_limit==0)
            $max_page=$num_max/$page_limit;
        else
            $max_page=floor($num_max/$page_limit)+1;
    }
    else
    {
        $min_page=$page-5;
        while($min_page<1)
            $min_page++;
        $max_page=$min_page+9;
        while((($max_page-1)*$page_limit>=$num_max)
            $max_page--;
        if($max_page-$min_page+1<10)
            $min_page=$max_page-9;
    }
    // var_dump($min_page);
    // var_dump($max_page);
}

// Jump to Page;
echo "<form id=\"form_jump_to_right_hand\" action=\"/EE101-Final_Project/Final_Project/
search.php#skip_multi">";
echo "<input class=\"input_button\" type=\"hidden\" name=\"key_word\" value=\"$key_word\">";
echo "Jump to: <input type=\"number\" name=\"page\" class=\"all_Turn_Page_jump_to_number\""
max=$page_MAX min=\"1\" required>&nbsp;&nbsp;";
echo "<input class=\"jump_page_button\" type=\"submit\" value=\"Go!\"></form>";

```

6.2.3 Advanced

In “advanced” .

For this version, JQuery and Ajax are implemented.

The features are just similar to that of V2.0. It will be introduced later, in section “Advanced Search” .

7 Search - Simple and Multi-field

In “search” .

The simple search V1.0 is just that of my Lab03. It’s simple and sometimes naïve. By the way, some widgets are designed.

For the V2.0, multi-field search is applied. Only one blank is shown. All the results, containing the target keywords in Fields Title, Authors_Name, and

Conference, will be displayed. These fields are of the same importance. Similarly, noinput will lead to an error.

Also refer to:

- (1) <https://blog.csdn.net/upxiaofeng/article/details/51460042>
- (2) https://blog.csdn.net/lies_joker/article/details/51684453

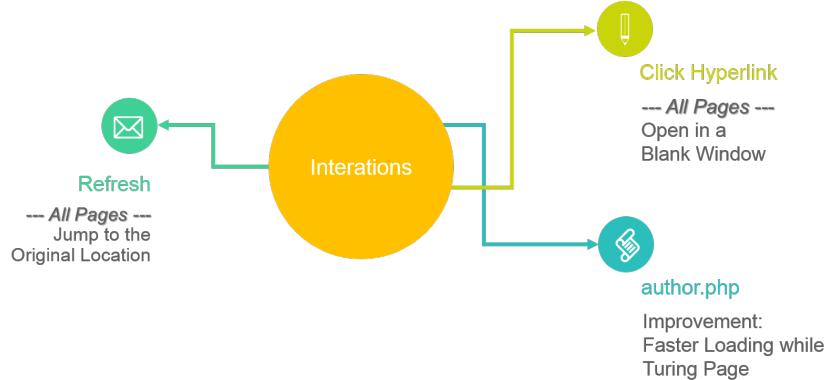
Effects:

Multi Field Search: jiawei han	
Title	Authors
ous information networks	jiawei han;
r growth methodology and implications	jiawei han; jian pei
n networks a structural analysis approach	yizhou sun; jiawei han;
used frequent graph patterns	xifeng yan; jiawei han
concept hierarchies for knowledge discovery in databases	jiawei han; yongjian fu;
ation retrieval a han character based approach	yuji matsumoto; md maruf hasan
nal information systems a preliminary design and experiment	osmar r zaiane; jiawei han;

Codes:

```
$url = "http://localhost:8983/solr/
lab02/select?indent=on&q=Title:".$query."^1+OR+Authors_Name:".$query.^0.7+OR+ConferenceName:".$query.^0.5&start=".($page_limit*($page-1))."&rows=".$page_limit."&wt=json&hl=on&hl.fl=Title,Authors_Name,ConferenceName&hl.simple.post=<%2Fb><%2Ffont>&hl.simple.pre=<font%20color%3D%23FF0000><b>";
```

8 Webpage Features



8.1 Refresh

For all the pages, after refreshing, the page will jump to the location where the user was previously browsing.

This function is implemented using cookies with reference to: <https://www.jb51.net/article/99749.htm>. Just include all the body parts in a *div onscroll=“SetH(this)”*.

Actually, to a certain extent, I have understood the codes. However, without any hints, I find it hard to do it all on my own.

Codes:

```
1  var _h = 0;
2  function SetH(o)
3  {
4      h = o.scrollTop
5      SetCookie("a", _h)
6  }
7
8  window.onload = function ()
9  { //页面加载时设置scrollTop高度
10    document.getElementById("x").scrollTop = GetCookie("a");
11  }
12
13  function SetCookie(sName, sValue)
14  {
15      document.cookie = sName + "=" + escape(sValue) + "; ";
16  }
17
18  function GetCookie(sName)
19  {
20      var aCookie = document.cookie.split(";");
21      for (var i = 0; i < aCookie.length; i++)
22      {
23          var aCrumb = aCookie[i].split("=");
24          if (sName == aCrumb[0])
25              return unescape(aCrumb[1]);
26      }
27      return 0;
28 }
```

8.2 Click Hyperlinks

Still, for all the pages, if a hyperlink is clicked, the new page will be opened in a new window.

It is easy to realize using parameter “target=_blank” .

Codes:

```
<a href="/EE101-Final_Project/Final_Project/index_search_advanced.php" target=_blank><button class="button" style="vertical-align:middle;display: inline;"><span>Advanced Searching</span></button>
```

8.3 Page Turning Anchors

It is **absolutely extremely disgusting** to be not guided to the head of the result table while turning pages!

Therefore, at the beginning of each result table, I set a anchor, which is added into the hyperlinks.

Codes:

```
echo "<a name=\"skip_multi\"></a>";
// .....
// Turn to the Next Page
echo "<td>";
$i=$page+1;
if (($i-1)*$page_limit<$num_max)
{
    echo "<a href=\"/EE101-Final_Project/Final_Project/
search.php?key_word=$key_word_temp&page=$i#skip_multi
\"><img src =\"/EE101-Final_Project/Final_Project/pics
/Turn_Page_empty.jpg\" id=\"search_Turn_Page_empty\"
></a>";
echo "</td><td>";
echo "<a class=\"output_href\" href=\"/EE101-Final_Project
/Final_Project/search.php?key_word=$key_word_temp
&page=$i#skip_multi\">Next</a>";
}
```

8.4 Loading Speed Improvements

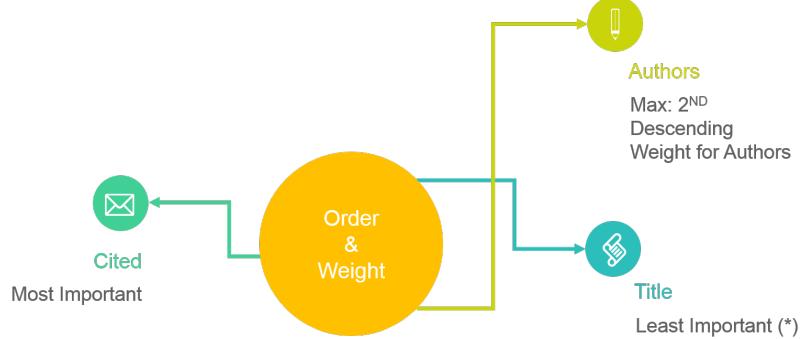
In addition, I improved the loading speed of the author page.

Originally, a request will be delivered to MySQL to search the most related affiliation. However, such a process is actually unnecessary at all!

So I pass the most related affiliation name while turning pages. The outcome is apparent!

9 Paper Recommendation

Ten recommended papers in academic formats, which can be hidden and shown, is available on the title page.



The order and weight account for three parameters. Cited times are given top priority, followed by authors and title. The authors of a certain paper are given a descending weight and the weight of the title is similar to that of the first author.

The implementation is similar to that of the multi-field search, except that in this case, different parameters are assigned varied weight.

Notice that:

- (1) The recommended papers can be clicked to acquire a deeper insight.
- (2) Few groups pay attention to the **academic format of paper information.**

Actually, the hide-and-show process is very very **fluent!**

To implement such function, I adopted JavaScript with reference to:

https://blog.csdn.net/baidu_35701759/article/details/76187236

It helps a lot at the beginning.

Effects:

Related Papers:

[1]. yizhou sun,jiawei han:mining heterogeneous information networks a structural analysis approach,SIGKDD,2013

[2]. marina danilevsky,jiawei han,ming jiranking based classification of heterogeneous information networks,SIGKDD,2011

[3]. jiawei han,yizhou sun,yintao yuranking based clustering of heterogeneous information networks with star network schema,SIGKDD,2009

Codes:

```
1  var btn = document.getElementById('btn')
2  var spread = document.getElementById('spread')
3  var iSpread = false
4  var height = spread.scrollHeight
5  var time = 420;
6  var interval = 8.4
7  var speed = height/(time/interval)
8
9  btn.onclick = function (e) {
10    btn.disabled = 'disabled'
11    if(!iSpread)
12    {
13      var speeds = 0
14      var timer = setInterval(function () {
15        speeds += speed
16        spread.style.height = speeds + 'px'
17
18        if(parseInt(spread.style.height) >=height){
19          clearTimeout(timer)
20          btn.disabled = ''
21        }
22      },interval)
23      this.innerHTML = 'Hide'
24    }
25  else
26  {
27    var speeds = height
28    this.innerHTML = 'Show'
29    var timer = setInterval(function () {
30      speeds -= speed
31      spread.style.height = speeds + 'px'
32      if(speeds <= 0)
33      {
34        clearTimeout(timer)
35        btn.disabled = ''
36      }
37    },interval)
38  }
39  iSpread = !iSpread
40 }
```

10 Advanced Search

10.1 Overview

The advanced search is key to a good search engine!



Here, we may reach some agreements on the names:

- (1) The area with the key word boxes is called: Search Bar;
- (2) The empty area now in yellow is called: Result Area.

This page is made separately using mainly Jquery. I was also in charge of **most of its front-end coding**.

For this part, since the unclarity and difficulty of division, I recommend that if you are interesting in the codes, please have a look at the source codes attached. I will point out the exact file names.

The main page (default file) is "index_search_advanced.php" .

10.2 Search Bar

Apparently, the “+” and “-” are bound to different incidents, which will send requests to “07_advanced_search_boxes.php” to update the number of boxes.

For the bool part, users can choose *And* or *Or*.

For the target part, users can choose *Title*, *Author* or *Conference*. For the input part, users can input anything. Phrases with spaces are also ok. There are some **details worth consideration**:

- (1) After clicking the add-or-delete buttons, the filled values must be kept. Therefore, while requesting, the url contains all the values of the existing items.

The approach is simple, just get the values according to the id and sequence them. Then, in the php file, GET these values and set it as the default values.

(2) When there is only one line, the “-” button must be hidden. Just change the display mode while updating the “innerHTML” content.

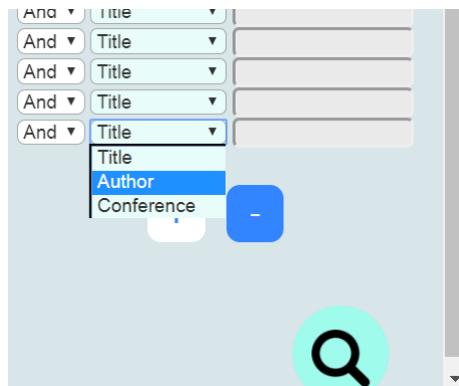
(3) Check for unfilled blanks and alert if exists when the search button is clicked. Loop works well. But attention must be paid to some small aspects.

(4) When there are too many lines, how should it appear? Actually, “overflow-y=auto” will do.

(5) Take care if there are spaces in the values! String replacement is required.

(6) Absolutely, for the first line, boolean selection is unnecessary. Thus, for both directions - add to two rows, delete to one row - deserves carefulness.

Effects:



10.3 Submit

After clicking the submit button, the parameters will be contained in the request, and the results will be shown in the Result Area.

It is a multi-field search. The keywords are given a slowly descending weight.

When clicked, a request containing all the parameters is sent to [“07_search_advanced.php”](#). The codes there are similar to that of “search” page.

Then, the results are loaded.

But the page turning area is loaded with the function “turn_page()”, where maximum and minimum page numbers are calculated and a request, containing the target page number, is sent to [“07_advanced_search_pages”](#). Please notice that, I am greatly anxious about the user-friendliness. Therefore, I made it

a rule that every successfull search will lead to the fact that the Search Bar **gradually slides and hides**. When it is hidden, a bar is for calling it back and viceversa. Of course, at first, this call-back-bar can not be seen. Here, the main part is similar to that of the paper recommendation hide-and-show of the “title” page, except for some details. You may refer to “07_show_hide.js” .

Effects:



10.4 Page Turning and Jump-to

Firstly, page turning codes in “07_advanced_search_pages.php” :

Similar to the traditional page turning functions, show the numbers and the images. But here, the images are bound with “`onclick()`” functions instead of hyperlinks. For the numbers, the hyperlink is bound with a “`onclick(submit_search($new_page,1))`” function and must not be jumped to “`href`” . Thus, this approach works:

```
href="#" onclick="submit_search($new_page,1); return false;"
```

Also refer to:

https://blog.csdn.net/fendou123_love/article/details/53585016

Secondly, “`submit_search(page,1)`” in “index_advanced_search.php” :

Turing pages are just the same as submitting a search request, except for the fact that the former requires the search bar always being kept hidden and

the latter being changed from visible to invisible. Therefore, I need a parameter to indicate such difference(1 for turning pages, 0 for submits).

Thirdly, jump-to function in “07_advanced_search_pages.php” :

The main codes are similar to those of the traditional search. But here, the JUMP button is bound with “`onclick=jump_to_submit()`” instead of a hyperlink.

Fourthly, “`jump_to_submit()`” in “index_advanced_search.php” :

The validation (i.e. filled/unfilled, maximum, minimum) of the inputted target page number is checked first. If it not valid, alert the error. Then, just the same as turning pages, call “`submit_search(new_page,1);`”

Effects:

Title	Authors	Conference
e business enterprise data mining	neal rothleder; usama fayyad; paul s bradley;	SIGKDD
mining time changing data streams	geoff hulten laurie spencer pedro domingos	SIGKDD
mining data records in web pages	robert grossman; bing liu; yanHong zhai;	SIGKDD
on the efficacy of data mining for security applications	ted e senator	SIGKDD
enriching textbooks through data mining	sreenivas gollapudi; krishnaram kenthapadi; raja p velu; rakesh agrawal; nitish srivastava;	WWW
dddm2007 domain driven data mining	graham j williams longbing cao philip s yu chengqi zhang yanchang zhao	SIGKDD
mining uncertain data with probabilistic guarantees	david w cheung; jieFeng cheng; reynold cheng; liwen sun;	SIGKDD
toward knowledge driven data mining	wanwick graco tatiana semenova eugene dubossarsky	SIGKDD
mining hepatitis data with temporal abstraction	katsuhiko takabayashi; truong dung nguyen; dung duc nguyen; tu bao ho; si quang le; saori kawasaki; hideto yokoi	SIGKDD
mining high speed data streams	geoff hulten pedro domingos	SIGKDD

Found 2771 results. Each page: 10 items. Alltogether: 277 pages.

Phanmunmununutom

Prev 1 2 3 4 5 6 7 8 9 10 Next

Jump to: Go!

11 Feedback

We value users' precious feedbacks. So we add a feedback page.

We are curious about four fields: the user's name, the e-mail address, the subject and what we can do for the user. Therefore, four blanks are required. The unfilled blanks will be alerted and the pending status will be clearly illustrated.



What's more, the log will be saved locally in a clear enough format.

==== UTC/GMT+08:00 2019-06-10 22:16:12 ===

Name: The Phantom
Email Address: ILoveIt@phantom.com
Subject: Say hello
Feedback: Hello World
***** END *****

Although PHPMailer seems to be much more convenient, something wrong occurred such as missing an authentic certificate. In the end, such approach was discarded. I take advantage of NodeJS and emailjs to implement the function.

Still, I mainly use JQeury to implement this section. When clicked, the valid messages will be connected together and passed to the requested .php file, where the cmd console will be run (using “exec()”) to call the nodejs mail-sending function.

Reference:

- (1) HTML - send e-mails: <http://www.fly63.com/article/detial/620>
- (2) PHP - exec(): https://blog.csdn.net/sinat_29862853/article/details/85253384
- (3) PHP - get system time: <https://www.jb51.net/article/148361.htm>
- (4) PHP - file read: <https://www.cnblogs.com/penghuwan/p/6884932.html>
- (5) PHP - file read: <https://www.cnblogs.com/penghuwan/p/6884932.html>
- (6) NodeJS - run with given parameters:
<https://www.jianshu.com/p/474e6d76f867>
<https://cloud.tencent.com/developer/article/1363526>

Codes:

```
xmlhttp.onreadystatechange=function()
{
    if (xmlhttp.readyState==4 && xmlhttp.status==200)
    {
        // document.getElementById("advanced_search_result_turn_page").innerHTML=xm
        lhttp.responseText;
        alert("Successful!");
        send_mail_btn.disabled = '';
        send_mail_btn.innerHTML="Submit";
    }
}

var name=$("#feedback_name")[0].value;
var address=$("#send_mail_address")[0].value;
var subject=$("#send_mail_subject")[0].value;
var message=$("#feedback_message")[0].value;
if(!name || !address || !message)
{
}
else
    send_mail_btn.innerHTML="Processing";
var url="/EE101-Final_Project/Final_Project/add-ons/07_mail_to.php?name="+name.
    replace(/ /g,"+")+"&address="+address+"&subject="+subject.replace(/ /g,"+").
    "&word="+message.replace(/ /g,"+");

console.log(url);
// request
xmlhttp.open("GET",url, true);
xmlhttp.send();

$name=$_GET["name"];
$address=$_GET["address"];
$subject=$_GET["subject"];
$word=$_GET["word"];

// acquire the time
date_default_timezone_set('PRC');
$time=date('Y-m-d H:i:s', time());

$log="== UTC/GMT+08:00 ".$time." ==\r\nName: ".$name."\r\nEmail Address: ".$address."
\r\nSubject: ".$subject."\r\nFeedback: ".$word."\r\n";
$log_end="***** END *****\r\n\r\n";
$DOCUMENT_ROOT = $_SERVER['DOCUMENT_ROOT'];
$fp = fopen("$DOCUMENT_ROOT/EE101-Final_Project/Final_Project/feedback/feedback.txt", 'a'
    );
// append content to file
fwrite($fp,$log.$log_end);
fclose($fp);

// var_dump($name);
// var_dump($address);
// var_dump($word);
// echo "node 07_mail_to.js \"\".$name."\\" \"$address.\" \"\".$word.\"\"";
print(exec('cd C:\xampp\htdocs\EE101-Final_Project\Final_Project\add-ons'));
exec("node 07_mail_to.js \"\".$name."\\" \"$address.\" \"UTC/GMT+08:00<br>".date("Y-m-d")."
\"");
```

[End] --- By Ziliang Guo 518030910273 ---

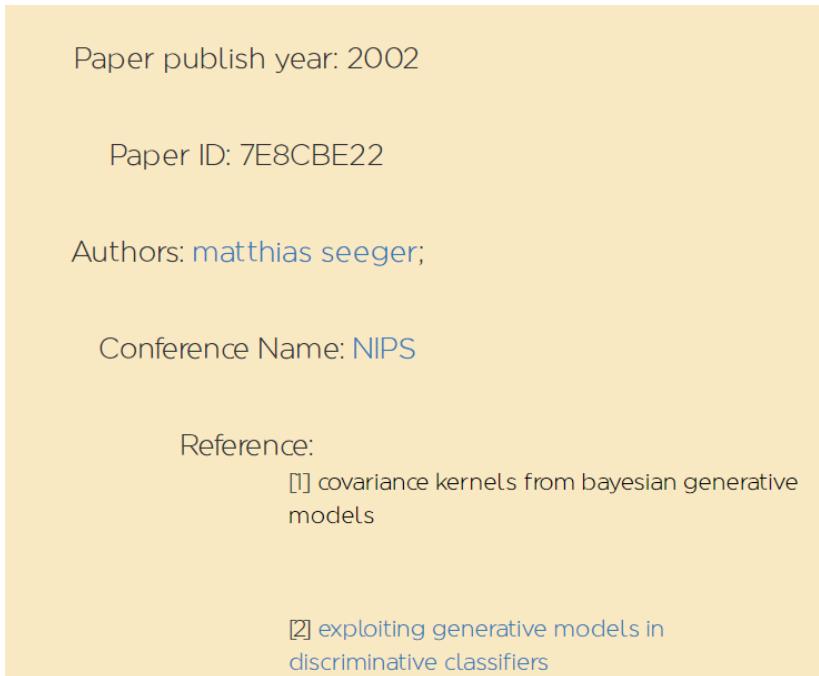
12 Hyperlinks

— By Aofan Jiang 518030910275 —

As required in the final project, I add the hyperlink of each title, author and conference. So the users can click to get more information about the result they want to search for.

12.1 Hyperlink of each title

Different from the table about the information of key words, the title page is simply a series of main information about the corresponding paper. It includes the paper ID, the publish year, the authors, conference name, and reference papers. This can make the whole page clearer, just as the following picture.



It might be ignored that the table of reference paper title of each paper is given in the original data. Although it is not required in the lab two, I still add it as a field in Solr to make sure that more detailed information can be gotten by the users. This explains the existence of references in the paper page.

What's more, to make a user-friendly website, you can see in the following picture that at the top of website, the link information about the title. It's also

clear and easy to find the information. It's linked by the symbol of addition.

localhost/EE101-Final_Project/Final_Project/title.php?title=learning+to+learn+with+the+informative+vector+machine

On the comparison, you can see even in Baidu NetDisk that the link is not so comfortable as ours. They are all connected with %2 while our website link are connected with + . Since all the information is imported in solr, the searching speed is quite fast with no delay.

=Auth%20Login%20Sucess&&bduss=&ssnerror=0&traceid=#/all?vmode=list&path=%2F我的资源%2F冰与火：权力的游戏1-6

12.2 Hyperlink of each conference

This page is also mainly an information table like the main searching page. It includes each paper's title, authors published on the given conference name. However, at the last column of table, the original conference name is replaced by the publish year of each paper published in this conference.

Title	Authors	Year
von mises fisher clustering models;	siddharth gopal; yiming yang;	2014
mixed membership matrix factorization;	lester mackey; michael i jordan; david j weiss	2010
a randomized anova procedure for comparing performance curves;	michael atighetchi; justus piater; xiaoqin zhang; paul r cohen;	1998
path normalcy analysis using nearest neighbor outlier detection	muthukumaran chandrasekaran; david luper; khaled rasheed; hamid r arabnia	2008
distributed stochastic gradient mcmc;	babak shahbaba; max welling; sungjin ahn;	2014
regularization of neural networks using dropconnect	li wan; xixin zhang; rob fergus; yann le cun; matthew d zeiler;	2013
automated cephalometric landmark localization using sparse shape and appearance models;	dirk vandermeulen; johannes keustermans; paul suetens; dirk smets;	2011
semi supervised learning through principal directions estimation	bernhard scholkopf; olivier chapelle; jason weston;	2003
forgetting counts constant memory inference for a dependent hierarchical pitman yor process;	nicholas bartlett; david pfaud; frank wood;	2010

12.3 others

Even at different pages among paper, author, conferences. Almost all the items are hyperlinked. So it means you can jump to different pages in any given page. Here comes an example of authors' information

Title	Authors	Conference
mining heterogeneous information networks;	jiawei han;	SIGKDD;
resource and knowledge discovery in global information systems a preliminary design and experiment	osmar r zaiane; jiawei han;	SIGKDD
mining frequent patterns by pattern growth methodology and implications;	jiawei han; jian pei;	SIGKDD;
mining heterogeneous information networks a structural analysis approach	yizhou sun; jiawei han;	SIGKDD
closegraph mining closed frequent graph patterns;	xifeng yan; jiawei han;	SIGKDD;
dynamic generation and refinement of concept hierarchies for knowledge discovery in databases	jiawei han; yongjian fu;	SIGKDD
on trivial solution and scale transfer problems in graph regularized nmf;	chris ding; jiawei han; quanquan gu;	IJCAI;

13 Code optimization

13.1 Solr

By testing, I find that the searching speed by solr is faster than mysql. So all the places that can be used in solr are changed from mysql.

13.2 Mysql

When using mysql, there are still some methods to improve the speed of program.

If the searching result is only one piece. For instance, the publish year, the affiliation and the conference and so on. We can limit the searching result numbers by add the code “LIMIT 1”. As a result, when mysql get one information about the result, it will stop the searching process, which is a great improvement in the speed of a program.

```
$result = mysqli_query($link, "SELECT AuthorName from authors where AuthorID='$author_id' limit 1");
```

Instead of using inner join, try to use more simple searching and output the result together is faster.

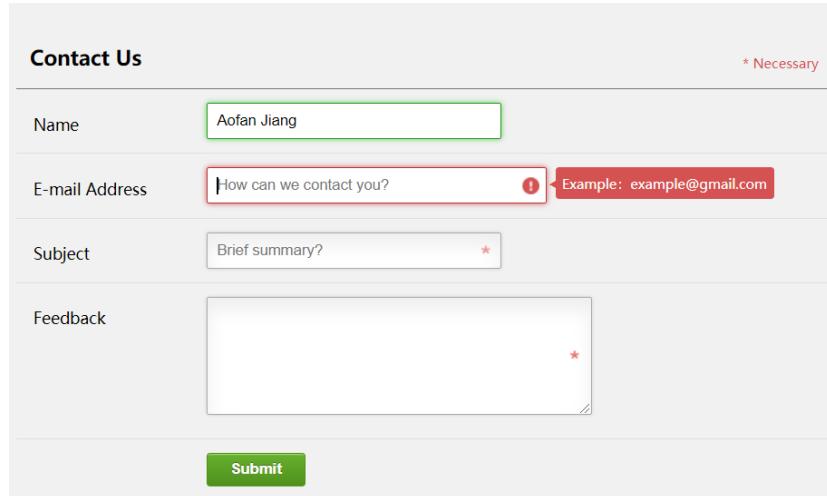
```
$affi_id_name_result = mysqli_query($link, "SELECT affiliations.AffiliationID, affiliations.
AffiliationName from (select AffiliationID, count(*) as cnt from paper_author_affiliation
where AuthorID='$author_id' and AffiliationID is not null group by AffiliationID order by cnt
desc) as tmp inner join affiliations on tmp.AffiliationID = affiliations.AffiliationID");
```

```
$affi_id_row=mysqli_fetch_row(mysqli_query($link,"SELECT AffiliationID, count(*) AS count FROM
paper_author_affiliation where AuthorID='$author_id' GROUP BY AffiliationID ORDER BY count DESC LIMIT 1"));
$affi_id=$affi_id_row[0];
$affi_id_name_result=mysqli_query($link,"SELECT AffiliationName from affiliations where AffiliationID='$
affi_id'");
```

As for the code in php, we can choose to use mysqli_fetch_row rather than mysqli_fetch_array since the information class is of each column is known to us.

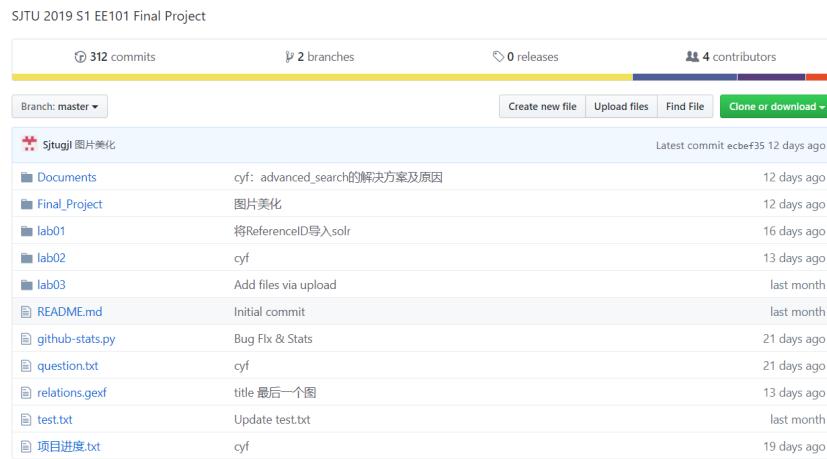
14 Contact and Open source

At the home page, you can find the word “contact us” . After clicking it, you will jump to a new page with a sophisticated chart. You can input your information and your information, we will get your feedback after you clicking the submit button.



The image shows a contact form titled "Contact Us". It includes fields for Name (Aofan Jiang), E-mail Address (How can we contact you? - placeholder, Example: example@gmail.com), Subject (Brief summary? - placeholder with red asterisk), and Feedback (a large text area with a red asterisk). A "Submit" button is at the bottom.

Also, you can find the word ” open source ” . After clicking it, you will jump to our project on Github website. At here, you can check all the detailed codes and corresponding documents.



The image shows a screenshot of a Github repository page for "SJTU 2019 S1 EE101 Final Project". The page displays statistics: 312 commits, 2 branches, 0 releases, and 4 contributors. It shows a list of files and their commit history:

File	Commit Message	Time
Sjtugji 图片美化	cyf: advanced_search的解决方案及原因	12 days ago
Documents	图片美化	12 days ago
Final_Project	将ReferenceID导入solr	16 days ago
lab01	cyf	13 days ago
lab02	Add files via upload	last month
lab03	Initial commit	last month
README.md	Bug Fix & Stats	21 days ago
github-stats.py	cyf	21 days ago
question.txt	title 最后一个图	13 days ago
relations.gexf	Update test.txt	last month
test.txt	cyf	19 days ago
项目进度.txt		