

编号_____

南京航空航天大学

毕 业 设 计

题 目 基于web大数据的用户交互关
 系可视化设计与开发

学生姓名

丛林

学 号

161310121

学 院

计算机科学与技术学院

专 业

计算机科学与技术

班 级

161310121

指导教师

夏正友 副教授

二〇一七年五月

南京航空航天大学

本科毕业设计（论文）诚信承诺书

本人郑重声明：所呈交的毕业设计（论文）（题目：基于 web 大数据的用户交互关系可视化设计与开发）是本人在导师的指导下独立进行研究所取得的成果。尽本人所知，除了毕业设计（论文）中特别加以标注引用的内容外，本毕业设计（论文）不包含任何其他个人或集体已经发表或撰写的成果作品。

作者签名:

年 月 日

(学号) :

基于 web 大数据的用户交互关系可视化设计与开发

摘 要

随着联网时代的来临，数据需要存储和处理量日益加大，庞大的数据和错综复杂的关系使对其存储、使用，展示方法成为了需要仔细研究考虑的问题，如何将大量数据有效存储，尽可能保留其关系，并将关系有效展示，针对类似情况其方法工具手段层出不穷。针对展示对象为独立个人的情况，将数据关系和内容已二维图标的方式传统而有效直观，配合相应工具，能将信息更好地传达给用户对象。

本论文使用了开源于互联网的数据可视化展示纯 Javascript 图标库工具 echarts，将用户关系数据用引力图的方式处理展示，提供了一直将图表展示图形化界面设置与前端的展示方式，纯 Javascript 的特性使其可以流畅地在 pc 或移动端设备进行展示查询，同时其高度可定制化的特性也使用户可在 web 端通过图表直接进行交互操作，提高了信息获取的效率，优化了数据展示的体验。

关键词：交互关系，可视化开发，echarts，数据展示

Visualized Design and Development of User Interaction Based on Web Large Data

Abstract

With the advent of the Internet era, the data needs to be stored and handled increasingly large, large data and intricate relationship to its storage, use, display method has become a need to carefully study the question, how to effectively store large amounts of data, May retain its relationship, and the relationship between the effective display, for a similar situation, its methods of tools endless. For the case of independent individuals, the data relationship and the content of the two-dimensional icon of the traditional way and effective and intuitive, with the appropriate tools, information can be better communicated to the user object.

This project uses the data from the Internet to visualize the display of pure Javascript icon library echarts, the user relationship data with gravitational map to deal with the show, the characteristics of pure Javascript so that it can be smooth in the pc or mobile device display query, while Its highly customizable features also allow users to interact directly through the graphs at the web end, improve the efficiency of information acquisition, and optimize the experience of data display.

Key Words: Interactive, Visual development, Echarts, Data display

摘 要.....	I
Abstract.....	II
第一章 引 言.....	1
1.1 课题研究的背景.....	1
1.2 课题研究的意义.....	1
1.3 国内外相关研究综述.....	2
第二章 用户关系可视化实现技术.....	5
2.1 java 技术.....	5
2.2 JSP 简介.....	8
2.3 JDBC 简介.....	9
2.4 MySQL 简介.....	10
2.5 Myeclipse 简介.....	11
2.6 echarts 简介.....	12
2.7 tomcat 简介.....	12
2.8 javascript 简介.....	13
2.9 ajax 简介.....	14
第三章 用户关系可视化的分析与设计.....	16
3.1 可视化需求分析.....	16
3.1.1 需求可发展条件分析.....	16
3.1.2 用户处理逻辑.....	17
3.1.3 可行性分析.....	17
3.2 开发方法.....	18
3.3 系统分析.....	18
3.3.1 系统功能结构.....	19
3.3.2 安全性分析.....	19
3.4 系统流程图.....	20
3.5 数据流程图.....	20

3.5 用户数据库构造.....	21
第四章 用户关系可视化的实现.....	23
4.1 模块功能实现介绍.....	23
4.1.1 引力图.....	23
4.1.2 环形图.....	25
4.2 细节算法与实现说明.....	26
4.2.1 后端细节设计.....	26
4.2.2 多级表搜索逻辑.....	28
4.2.3 前端页面细节设计.....	30
第五章 总结与展望.....	35
5.1 总结.....	35
5.2 展望.....	36
参 考 文 献.....	37
致 谢.....	38

第一章 引言

1.1 课题研究的背景

社交网络为以互联网为依托为人群提供的交流服务，其形式多种多样，电邮，论坛社区等等均在其列。

多数社交网络会提供多种让用户交互起来的方式，可以为聊天、寄信、影音、文件分享、博客、新闻组等^[1]。

社交网络为信息的交流与分享提供了新的途径，近年来随着中国互联网的发展，社交网络已经完全渗透入生活中，成为工作学习娱乐不可缺少的一部分。目前中国开发性社交网络中最大也最具代表意义的是新浪微博。新浪微博中 用户关系并非是静态的、片面的，而是不断地随着消息的流转 而动态发生改变。这些消息的流转、用户关系的动态变化，又能对现实社会产生潜移默化的反作用力^[2]。

在人类活动的许多领域，特别是随着互联网的发展，海量的信息内容被公开提供给分析者，大规模数据的分析和应用成为了值得挑战的难题，其中数据的可视化分析因其直观的信息内容表达方式而备受青睐。

可视化分析是由多学科交叉发展而来的新兴的学科领域，随着图形化界面以及交互技术的发展，许多精力都集中在信息可视化领域的研究中，同时，越来越多的研究者认识到，通过有效的可视化技术让用户参与到知识发现和数据挖掘的过程中有巨大潜力^[3]。

1.2 课题研究的意义

海量和社交网络用户和其及其强大的活跃度为互联网提供了巨大丰富真实可供参考的数据，如何对这些数据进行收集，分析，并获得有价值信息成为了许多互联网从业公司或研究机构关注的焦点。

通过大数据分析技术对所获数据进行分析，将所得结论可视化展示，这种方式能直观地提供可供理解的数据信息，客观形象描述大规模数据分析结果，为相关分析者，策略定制者或行业入门了解者提供参考。

1.3 国内外相关研究综述

社交网络发展势头迅猛，可供分析数据多样丰富，sun 等研究过国外社交网络 facebook，并基于其用户建立了社交网络分析系统，围绕 Facebook 上不同种类的传播社区间的关系，发现并确定其融合趋势^[4]。

Cha 等对 Flickr 数据进行探索，构造了照片和拍照人之间的关系，并探索了信息传播的范围和社交网络中的口碑传播内容^[5]。

Sakaki 等在基于 Twitter 数据上提出了一个实时时间检测系统。他们将 Twitter 用户视为事件的实时检测者并进行追踪^[6]。

Ho 等将注意力放在衡量信息是如何传播的，考虑怎样莲花社交网络上的用户传播信息的能力以及衡量传播的范围等。他们基于信息覆盖人数、信息传播速度、传播覆盖的地理范围。其系统显示了传播的路径以及相关网络图^[7]。

国内方面依托于新浪微博的活跃，陈胜楠探索了咱新浪微博中自动化数据分析挖掘系统的设计和相关技术的思路，对于社交网络中的用户生成文本内容进行分析，提高了社交平台采集功能的适应性。

陈召群等针对社交网络中的凝聚子群提出了层次化的可视化方法。另外一些研究者针对特定问题，使用可视化工具如 Gephi、Pajek、Cytoscape 和 SocNetV 等进行分析，提出了基于话题的社交网络分析技术，并使用可视化工具给出了可视化结果。针对大规模社交网络可视化的研究较少，提出合并邻接矩阵相邻的元素形成“蜂巢”，能够对大规模网络进行可视化，但是仅对邻接矩阵进行合并处理，无法很好地保持原网络的结构特征。总得来说，大规模社交网络可视化的难点在于处理海量数据的同时需要保持原网络的结构特点。提出了一种基于谱聚类算法的大规模社交网络可视化方法。该方法的创新点在于利用社交网络中幂律分布的特点，将对网络结构更重要的核心用户从普通用户中区分出来，能够在保持社交网络结构特点的同时将千万量级的用户展示在二维可视化平面中。本文提出的方法对在线社交网络分析有重要的意义，并且能够扩展到其它复杂网络中^[8]。

丁晓东利用国外社交网络分析领域的成熟研究成果，在一个已有的网络监管系统中提出并实现一个用于分析虚拟网络上对象行为和虚拟社交网络关系的社交网络分析工具，并用此实现了用于网络监管的社交网络可视化分析工具^[9]。

陈玉英将机器学习、数据挖掘的相关理论、技术手段应用于微博，并创新性地对这种社交网络中信息传播及影响模型进行了改进，使模型对信息传播效果得到了进一步提升，其从用户行为特征，微博特征，关系特征进行分析，提出了人类动力学模型的改良^[10]

在知识可视化的理论与方法上，赵国庆等人提供了详细的知识可视化的各种方案，其中动态图表可视化的方法有极高的信息传递效率，并得出其实未来可视化领域发展趋势的观点^[11]。

针对社交网络中的数据，往往搜索结论复杂庞大，李绪等在社交网络数据个性化推荐的可视化方法中也做了极多尝试。针对大规模社交网络应用中检索结果过于庞大复杂的问题，将个性化推荐与可视化相结合，用于在大量数据中找到用户感兴趣的信息。在开拓网络缩放算法的基础上，提出关键信息显示算法，能够区别显示社交网络关系图中用户相对重要的信息和次要信息，增强关联度较高数据的显示效果。将带权值的力导向布局算法应用于用户关系聚类中，通过在二维显示空间中合理安排节点布局，达到减少用户认知负担和个性化推荐的目的。设计并实现个性化推荐的可视化工具 HRVis，在 Movielens 数据集上进行测试，结果表明，HRVis 能够强调显示具有良好社会关系的重要用户以及与用户相似的关联用户，获得较好的可视推荐效果^[12]。

面对社交网络中多维数据，通过有效的可视化技术，用户参与到知识发现和数据挖掘的过程中。李权等人着重研究将可视分析技术用于多维数据规律发现、多维数据聚类可视分析及结合典型的拥有多维数据特征的热门流行的微博社交网络进行应用。在分析多维数据时，其提出一种新的可视呈现方法，用于揭示数据中隐藏的规律、模式及数据关联性。传统分析多维数据方法如平行坐标，虽然能对数据整体模式进行直观呈现，但与由维度映射成轴的摆放直接相关，且不易观察某维度上数值变化是如何引起其他维度数值、模式等信息改变。其分别从数据及可视呈现对轴摆放进行深入研究，并从处理后的数据中抽取模式信息来指导数据重排列，布局方式是由一系列自动计算及交互调整组合生成的，将数值、模式及梯度等进行集成，映射为三层同心圆组，提供全局重排或局部重排等交互操作。最后，通过分析若干数据集验证本方法的有效性。其设计并实现一个在可视交互环境中的多维数据聚类分析框架。在聚类生成阶段，将

多维数据映射为图并使用力导向布局展示其内部结构,使用马尔可夫图聚类给用户一个对数据的初始了解。使用 Sammon 投影将多维数据映射至二维平面,并支持手工选择聚类中心,还集成了 k-means 聚类算法。聚类评估阶段,通过平行聚类视图进行聚类分析,观察核心数据,调整维度权重,排除噪声,并及时将用户操作结果迭代反馈给所训练的数据模型。该方法能有效解决传统聚类算法的盲目性,充分利用专家领域知识,提高聚类结果可信度。 微博数据是典型的多维数据集。追踪微博转发传播是一项重要有意义的工作。基于可视分析角度,对传播转发的高维属性,提出三种新的策略: 层次动态布局,累计转发简化布局及微群布局。对动态演变过程,提出两种新的模型: 基于层次变化及基于时间变化的数据模型。提供了一系列交互功能,帮助用户动态理解探索微博数据。使用新浪微博数据验证系统的高效性及有效性,并在新浪微博得到了广泛内部应用,极大改善了数据分析师的工作效率。

第二章 用户关系可视化实现技术

2.1 java 技术

java 语言于 1995 年正式发布，其有面向对象、分布式等特点，同时也是我国企业开发最常用语言之一，其为基于网络计算的开发语言，因为其的平台无关性，使得其被广泛应用与大量使用，本论文中所实现的用户关系可视化系统因此选用 java 语言作为底层实现基本逻辑的工具，在具体使用过程中起灵活性令人赞叹，也因用户厂商对其的广泛支持另项目部署工作变得极其简单通畅^[13]。

(1)简单性

Java 看起来设计得很像 C++，但是为了使语言小和容易熟悉，设计者们把 C++语言中许多可用的特征去掉了，这些特征是一般程序员很少使用的。例如，Java 不支持 go to 语句，代之以提供 break 和 continue 语句以及异常处理。Java 还剔除了 C++的操作符过载（overload）和多继承特征，并且不使用主文件，免去了预处理程序。因为 Java 没有结构，数组和串都是对象，所以不需要指针。Java 能够自动处理对象的引用和间接引用，实现自动的无用单元收集，使用户不必为存储管理问题烦恼，能更多的时间和精力花在研发上。

(2)面向对象

Java 是一个面向对象的语言。对程序员来说，这意味着要注意应中的数据 and 操纵数据的方法（method），而不是严格地用过程来思考。在一个面向对象的系统中，类（class）是数据和操作数据的方法的集合。数据和方法一起描述对象（object）的状态和行为。每一对象是其状态和行为的封装。类是按一定体系和层次安排的，使得子类可以从超类继承行为。在这个类层次体系中有一个根类，它是具有一般行为的类。Java 程序是用类来组织的。

Java 还包括一个类的扩展集合，分别组成各种程序包（Package），用户可以在自己的程序中使用。例如，Java 提供产生图形用户接口部件的类（java.awt 包），这里 awt 是抽象窗口工具集（abstract windowing toolkit）的缩写，处理输入输出的类（java.io 包）和支持网络功能的类（java.net 包）。

(3)分布性

Java 设计成支持在网络上应用，它是分布式语言。Java 既支持各种层次的网络连接，又以 Socket 类支持可靠的流（stream）网络连接，所以用户可以产生分布式的客户机和服务器。网络变成软件应用的分布运载工具。Java 程序只要编写一次，就可到处运行。

(4)编译和解释性

Java 编译程序生成字节码（byte-code），而不是通常的机器码。Java 字节码提供对体系中性的目标文件格式，代码设计成可有效地传送程序到多个平台。Java 程序可以在任何实现了 Java 解释程序和运行系统（run-time system）的系统上运行。

在一个解释性的环境中，程序开发的标准“链接”阶段大大消失了。如果说 Java 还有一个链接阶段，它只是把新类装进环境的过程，它是增量式的、轻量级的过程。因此，Java 支持快速原型和容易试验，它将导致快速程序开发。这是一个与传统的、耗时的“编译、链接和测试”形成鲜明对比的精巧的开发过程。

(5)稳健性

Java 原来是用作编写消费类家用电子产品软件的语言，所以它是被设计成写高可靠和稳健软件的。Java 消除了某些编程错误，使得用它写可靠软件相当容易。

Java 是一个强类型语言，它允许扩展编译时检查潜在类型不匹配问题的功能。Java 要求显式的方法声明，它不支持 C 风格的隐式声明。这些严格的要求保证编译程序能捕捉调用错误，这就导致更可靠的程序。

可靠性方面最重要的增强之一是 Java 的存储模型。Java 不支持指针，它消除重写存储和讹误数据的可能性。类似地，Java 自动的“无用单元收集”预防存储漏泄和其它有关动态存储分配和解除分配的有害错误。Java 解释程序也执行许多运行时的检查，诸如验证所有数组和串访问是否在界限之内。

异常处理是 Java 中使得程序更稳健的另一个特征。异常是某种类似于错误的异常条件出现的信号。使用 try/catch/finally 语句，程序员可以找到出错的处理代码，这就简化了出错处理和恢复的任务。

(6)安全性

Java 的存储分配模型是它防御恶意代码的主要方法之一。Java 没有指针，所以程序员不能得到隐蔽起来的内幕和伪造指针去指向存储器。更重要的是，Java 编译程序不处理存储安

排决策，所以程序员不能通过查看声明去猜测类的实际存储安排。编译的 Java 代码中的存储引用在运行时由 Java 解释程序决定实际存储地址。

Java 运行系统使用字节码验证过程来保证装载到网络上的代码不违背任何 Java 语言限制。这个安全机制部分包括类如何从网上装载。例如，装载的类是放在分开的名字空间而不是局部类，预防恶意的小应用程序用它自己的版本来代替标准 Java 类。

(7)可移植性

Java 使得语言声明不依赖于实现的方面。例如，Java 显式说明每个基本数据类型的大小和它的运算行为（这些数据类型由 Java 语法描述）。

Java 环境本身对新的硬件平台和操作系统是可移植的。Java 编译程序也用 Java 编写，而 Java 运行系统用 ANSIC 语言编写。

(8)高性能

Java 是一种先编译后解释的语言，所以它不如全编译性语言快。但是有些情况下性能是很要紧的，为了支持这些情况，Java 设计者制作了“及时”编译程序，它能在运行时把 Java 字节码翻译成特定 CPU（中央处理器）的机器代码，也就是实现全编译了。

Java 字节码格式设计时考虑到这些“及时”编译程序的需要，所以生成机器代码的过程相当简单，它能产生相当好的代码。

(9)多线索性

Java 是多线索语言，它提供支持多线索的执行（也称为轻便过程），能处理不同任务，使具有线索的程序设计很容易。Java 的 lang 包提供一个 Thread 类，它支持开始线索、运行线索、停止线索和检查线索状态的方法。

Java 的线索支持也包括一组同步原语。这些原语是基于监督程序和条件变量风范，由 C.A.R.Haore 开发的广泛使用的同步化方案。用关键词 synchronized，程序员可以说明某些方法在一个类中不能并发地运行。这些方法在监督程序控制之下，确保变量维持在一个一致的状态。

(10)动态性

Java 语言设计成适应于变化的环境，它是一个动态的语言。例如，Java 中的类是根据需要载入的，甚至有些是通过网络获取的^[14]。

2.2 JSP 简介

JSP 为 java 服务器页面，其为一种由多家软件公司共同建立的技术标准，通过 JSP 技术，使得 java 代码能直接利用特定标签嵌入 jsp 页面中，实现了仅有 jsp 页面同 servlet 进行通讯的功能，本论文所实现功能考虑到后期需和整体系统进行对接并且未采用多人协作的版本控制系统，因此将所有 java 逻辑代码写入 jsp 页面中，虽然简单粗暴，但确实达到了良好的效果，使得移植转移工作变得极其简单方便。JSP 全名为 Java Server Pages，中文名叫 java 服务器页面，其根本是一个简化的 Servlet 设计，它[1] 是由 Sun Microsystems 公司倡导、许多公司参与一起建立的一种动态网页技术标准。JSP 技术有点类似 ASP 技术，它是在传统的网页 HTML（标准通用标记语言的子集）文件(*.htm,*.html)中插入 Java 程序段(Scriptlet)和 JSP 标记(tag)，从而形成 JSP 文件，后缀名为(*.jsp)。用 JSP 开发的 Web 应用是跨平台的，既能在 Linux 下运行，也能在其他操作系统上运行^[15]。

它实现了 Html 语法中的 java 扩展（以 <%, %>形式）。JSP 与 Servlet 一样，是在服务器端执行的。通常返回给客户端的就是一个 HTML 文本，因此客户端只要有浏览器就能浏览。

JSP 技术使用 Java 编程语言编写类 XML 的 tags 和 scriptlets，来封装产生动态网页的处理逻辑。网页还能通过 tags 和 scriptlets 访问存在于服务端的资源的应用逻辑。JSP 将网页逻辑与网页设计的显示分离，支持可重用的基于组件的设计，使基于 Web 的应用程序的开发变得迅速和容易。JSP(JavaServer Pages)是一种动态页面技术，它的主要目的是将表示逻辑从 Servlet 中分离出来。

Java Servlet 是 JSP 的技术基础，而且大型的 Web 应用程序的开发需要 Java Servlet 和 JSP 配合才能完成。JSP 具备了 Java 技术的简单易用，完全的面向对象，具有平台无关性且安全可靠，主要面向因特网的所有特点。

JSP2.0 中的一个主要功能是 JSP fragment，它的基本特点是可以使处理 JSP 的容器推迟评估 JSP 标记属性。我们知道一般 JSP 是首先评估 JSP 标记的属性，然后在处理 JSP 标记时使用这些属性，而 JSP fragment 提供了动态的属性。也就是说，这些属性在 JSP 处理其标记体时是可以被改变的。JSP 需要将这样的属性定义为 javax.servlet.jsp.tagext.JspFragment 类型。当 JSP 标记设置成这种形式时，这种标记属性实际上的处理方法类似于标记体。在实现标记

的程序中，标记属性可以被反复评估多次。这种用法称为 JSP fragment。JSP fragment 还可以定义在一个 SimpleTag 处理程序中使用的自制标记动作。像前面例子说明的，getJspBody 返回一个 JspFragment 对象并可以在 doTag 方法中多次使用。需要注意的是，使用 JSP fragment 的 JSP 只能有一般的文本和 JSP action，不能有 scriptlet 和 scriptlet 表达式。

2.3 JDBC 简介

JDBC 为一整套使 java 运行 sql 语句的 API，本项目使用纯 JDBC+直接写入 sql 语句对数据库进行操作，考虑到系统未来方便对接并未采用目前流行的 java 常用数据链接框架

（mybatis，hibernate），对于 sql 语句功能简单的需求可快速实现，缺点为稳定性可拓展性相对来说较差。有了 JDBC，向各种关系数据发送 SQL 语句就是一件很容易的事。换言之，有了 JDBC API，就不必为访问 Sybase 数据库专门写一个程序，为访问 Oracle 数据库又专门写一个程序，或为访问 Informix 数据库又编写另一个程序等等，程序员只需用 JDBC API 写一个程序就够了，它可向相应数据库发送 SQL 调用。同时，将 Java 语言和 JDBC 结合起来使程序员不必为不同的平台编写不同的应用程序，只须写一遍程序就可以让它在任何平台上运行，这也是 Java 语言“编写一次，处处运行”的优势。

Java 数据库连接体系结构是用于 Java 应用程序连接数据库的标准方法。JDBC 对 Java 程序员而言是 API，对实现与数据库连接的服务提供商而言是接口模型。作为 API，JDBC 为程序开发提供标准的接口，并为数据库厂商及第三方中间件厂商实现与数据库的连接提供了标准方法。JDBC 使用已有的 SQL 标准并支持与其它数据库连接标准，如 ODBC 之间的桥接。JDBC 实现了所有这些面向标准的目标并且具有简单、严格类型定义且高性能实现的接口。

Java 具有坚固、安全、易于使用、易于理解和可从网络上自动下载等特性，是编写数据库应用程序的杰出语言。所需要的只是 Java 应用程序与各种不同数据库之间进行对话的方法。而 JDBC 正是作为此种用途的机制。

JDBC 扩展了 Java 的功能。例如，用 Java 和 JDBC API 可以发布含有 applet 的网页，而该 applet 使用的信息可能来自远程数据库。企业也可以用 JDBC 通过 Intranet 将所有职员连到一个或多个内部数据库中（即使这些职员所用的计算机有 Windows、Macintosh 和 UNIX 等各种不同的操作系统）。随着越来越多的程序员开始使用 Java 编程语言，对从 Java 中便捷地访问数据库的要求也在日益增加。

MIS 管理员们都喜欢 Java 和 JDBC 的结合，因为它使信息传播变得容易和经济。企业可继续使用它们安装好的数据库，并能便捷地存取信息，即使这些信息是储存在不同数据库管理系统上。新程序的开发期很短。安装和版本控制将大为简化。程序员可只编写一遍应用程序或只更新一次，然后将它放到服务器上，随后任何人就都可得到最新版本的应用程序。对于商务上的销售信息服务，Java 和 JDBC 可为外部客户提供获取信息更新的更好方法。

2.4 MySQL 简介

本功能选用关系型数据库软件 MySQL 作为数据库系统存储用户数据资料，从互联网通过爬虫等抓取的非关系形数据通过转换为关系型数据使用 MySQL 进行存储，功能从数据库中直接读取数据与关系构成可视化图表进行显示，同时 MySQL 相比同类数据库系统其轻量跨平台的优势也使得开发过程减少了大量的工作量，使得大量开发工作都能在宿舍或地铁零散时间中完成^[16]。

MySQL 是一个关系型数据库管理系统，由瑞典 MySQL AB 公司开发，目前属于 Oracle 旗下产品。MySQL 是最流行的关系型数据库管理系统之一，在 WEB 应用方面，MySQL 是最好的 RDBMS (Relational Database Management System，关系数据库管理系统) 应用软件。

MySQL 是一种关系数据库管理系统，关系数据库将数据保存在不同的表中，而不是将所有数据放在一个大仓库内，这样就增加了速度并提高了灵活性。

MySQL 所使用的 SQL 语言是用于访问数据库的最常用标准化语言。MySQL 软件采用了双授权政策，分为社区版和商业版，由于其体积小、速度快、总体拥有成本低，尤其是开放源码这一特点，一般中小型网站的开发都选择 MySQL 作为网站数据库。

由于其社区版的性能卓越，搭配 PHP 和 Apache 可组成良好的开发环境。

MyISAM MySQL 5.0 之前的默认数据库引擎，最为常用。拥有较高的插入，查询速度，但不支持事务

InnoDB 事务型数据库的首选引擎，支持 ACID 事务，支持行级锁定, MySQL 5.5 起成为默认数据库引擎

BDB 源自 Berkeley DB，事务型数据库的另一种选择，支持 Commit 和 Rollback 等其他事务特性

Memory 所有数据置于内存的存储引擎，拥有极高的插入，更新和查询效率。但是会占用和数据量成正比的内存空间。并且其内容会在 MySQL 重新启动时丢失

Merge 将一定数量的 MyISAM 表联合而成一个整体，在超大规模数据存储时很有用

Archive 非常适合存储大量的独立的，作为历史记录的数据。因为它们不经常被读取。

Archive 拥有高效的插入速度，但其对查询的支持相对较差

Federated 将不同的 MySQL 服务器联合起来，逻辑上组成一个完整的数据库。非常适合分布式应用^[17]。

Cluster/NDB 高冗余的存储引擎，用多台数据机器联合提供服务以提高整体性能和安全性。适合数据量大，安全和性能要求高的应用

CSV：逻辑上由逗号分割数据的存储引擎。它会在数据库子目录里为每个数据表创建一个 .csv 文件。这是一种普通文本文件，每个数据行占用一个文本行。CSV 存储引擎不支持索引。

BlackHole：黑洞引擎，写入的任何数据都会消失，一般用于记录 binlog 做复制的中继

EXAMPLE 存储引擎是一个不做任何事情的空引擎。它的目的是作为 MySQL 源代码中的一个例子，用来演示如何开始编写一个新存储引擎。同样，它的主要兴趣是对开发者。**EXAMPLE** 存储引擎不支持索引。

另外，MySQL 的存储引擎接口定义良好。有兴趣的开发者可以通过阅读文档编写自己的存储引擎。

2.5 Myeclipse 简介

项目所用的开发集成环境为 Myeclipse，其为在 eclipse 的基础上加入常用插件与更多针对性 web 开发工具的企业级开发环境，因为本项目为纯 jsp 实现，所以选用其他 IDE 同样可以很好地完成。

MyEclipse 企业级工作平台（MyEclipseEnterprise Workbench，简称 MyEclipse）是对 EclipseIDE 的扩展，利用它我们可以在数据库和 JavaEE 的开发、发布以及应用程序服务器的整合方面极大的提高工作效率。它是功能丰富的 JavaEE 集成开发环境，包括了完备的编码、调试、测试和发布功能，完整支持 HTML, Struts, JSP, CSS, Javascript, Spring, SQL, Hibernate。

MyEclipse 是一个十分优秀的用于开发 Java, J2EE 的 Eclipse 插件集合, MyEclipse 的功能非常强大, 支持也十分广泛, 尤其是对各种开源产品的支持十分不错。MyEclipse 可以支持 Java Servlet, AJAX, JSP, JSF, Struts, Spring, Hibernate, EJB3, JDBC 数据库链接工具等多项功能。可以说 MyEclipse 是几乎囊括了目前所有主流开源产品的专属 eclipse 开发工具。

2.6 echarts 简介

echarts 为国内百度支持的开源纯 Javascript 图表库, 也是本可视化功能实现的核心工具, 其简便单纯, 提供的引力图和圆环布局完美地切合数据可视化需求, 同时支持图标缩放, 扁平化设计也为展示图标提供了更进一步易读性和可操作性

ECharts, 一个纯 Javascript 的图表库, 可以流畅的运行在 PC 和移动设备上, 兼容当前绝大部分浏览器 (IE8/9/10/11, Chrome, Firefox, Safari 等), 底层依赖轻量级的 Canvas 类库 ZRender, 提供直观, 生动, 可交互, 可高度个性化定制的数据可视化图表。

ECharts 3 中更是加入了更多丰富的交互功能以及更多的可视化效果, 并且对移动端做了深度的优化。

ECharts 提供了常规的折线图, 柱状图, 散点图, 饼图, K 线图, 用于统计的盒形图, 用于地理数据可视化的地图, 热力图, 线图, 用于关系数据可视化的关系图, treemap, 多维数据可视化的平行坐标, 还有用于 BI 的漏斗图, 仪表盘, 并且支持图与图之间的混搭。

你可以在下载界面下载包含所有图表的构建文件, 如果只是需要其中一两个图表, 又嫌包含所有图表的构建文件太大, 也可以在在线构建中选择需要的图表类型后自定义构建。

2.7 tomcat 简介

tomcat 为目前最先进且流行的开源 web 服务器, 由于其开发过程中 SUN 公司的参与, 其总能完美体现 serclet 与 jsp 规范, 免费开源且先进流行被广泛支持, 使其为个人 web 应用开发服务器的不二之选。

Tomcat 是 Apache 软件基金会 (Apache Software Foundation) 的 Jakarta 项目中的一个核心项目, 由 Apache、Sun 和其他一些公司及个人共同开发而成。由于有了 Sun 的参与和支持, 最新的 Servlet 和 JSP 规范总是能在 Tomcat 中得到体现, Tomcat 5 支持最新的 Servlet 2.4

和 JSP 2.0 规范。因为 Tomcat 技术先进、性能稳定，而且免费，因而深受 Java 爱好者的喜爱并得到了部分软件开发商的认可，成为目前比较流行的 Web 应用服务器。

Tomcat 服务器是一个免费的开放源代码的 Web 应用服务器，属于轻量级应用服务器，在中小型系统和并发访问用户不是很多的场合下被普遍使用，是开发和调试 JSP 程序的首选。对于一个初学者来说，可以这样认为，当在一台机器上配置好 Apache 服务器，可利用它响应 HTML（标准通用标记语言下的一个应用）页面的访问请求。实际上 Tomcat 部分是 Apache 服务器的扩展，但它是独立运行的，所以当你运行 tomcat 时，它实际上作为一个与 Apache 独立的进程单独运行的。

诀窍是，当配置正确时，Apache 为 HTML 页面服务，而 Tomcat 实际上运行 JSP 页面和 Servlet。另外，Tomcat 和 IIS 等 Web 服务器一样，具有处理 HTML 页面的功能，另外它还是一个 Servlet 和 JSP 容器，独立的 Servlet 容器是 Tomcat 的默认模式。不过，Tomcat 处理静态 HTML 的能力不如 Apache 服务器。目前 Tomcat 最新版本为 9.0。

2.8 javascript 简介

JavaScript 一种直译式脚本语言，是一种动态类型、弱类型、基于原型的语言，内置支持类型。它的解释器被称为 JavaScript 引擎，为浏览器的一部分，广泛用于客户端的脚本语言，最早是在 HTML（标准通用标记语言下的一个应用）网页上使用，用来给 HTML 网页增加动态功能。

在 1995 年时，由 Netscape 公司的 Brendan Eich，在网景导航者浏览器上首次设计实现而成。因为 Netscape 与 Sun 合作，Netscape 管理层希望它外观看起来像 Java，因此取名为 JavaScript。但实际上它的语法风格与 Self 及 Scheme 较为接近。

为了取得技术优势，微软推出了 JScript，CEnvi 推出 ScriptEase，与 JavaScript 同样可在浏览器上运行。为了统一规格，因为 JavaScript 兼容于 ECMA 标准，因此也称为 ECMAScript。

JavaScript 是一种属于网络的脚本语言，已经被广泛用于 Web 应用开发，常用来为网页添加各式各样的动态功能，为用户提供更流畅美观的浏览效果。通常 JavaScript 脚本是通过嵌入在 HTML 中来实现自身的功能的。

- 是一种解释性脚本语言（代码不进行预编译）。

- 主要用来向 HTML（标准通用标记语言下的一个应用）页面添加交互行为。
- 可以直接嵌入 HTML 页面，但写成单独的 js 文件有利于结构和行为的分离。
- 跨平台特性，在绝大多数浏览器的支持下，可以在多种平台下运行（如 Windows、Linux、Mac、Android、iOS 等）。

Javascript 脚本语言同其他语言一样，有它自身的基本数据类型，表达式和算术运算符及程序的基本程序框架。Javascript 提供了四种基本的数据类型和两种特殊数据类型用来处理数据和文字。而变量提供存放信息的地方，表达式则可以完成较复杂的信息处理^[19]。

2.9 ajax 简介

AJAX 不是一种新的编程语言，而是一种用于创建更好更快以及交互性更强的 Web 应用程序的技术。

使用 Javascript 向服务器提出请求并处理响应而不阻塞用户！核心对象 XMLHttpRequest。通过这个对象，您的 JavaScript 可在不重载页面的情况与 Web 服务器交换数据，即在不需刷新页面的情况下，就可以产生局部刷新的效果。

AJAX 在浏览器与 Web 服务器之间使用异步数据传输（HTTP 请求），这样就可使网页从服务器请求少量的信息，而不是整个页面。

AJAX 可使因特网应用程序更小、更快，更友好。

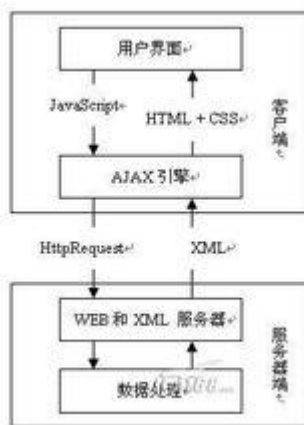


图 2.1 ajax 工作原理

AJAX 是一种独立于 Web 服务器软件的浏览器技术。AJAX 基于下列 Web 标准：JavaScriptXMLHTMLCSS 在 AJAX 中使用的 Web 标准已被良好定义，并被所有的主流浏览器支持。AJAX 应用程序独立于浏览器和平台。

Web 应用程序较桌面应用程序有诸多优势：它们能够涉及广大的用户，它们更易安装及维护，也更易开发。

不过，因特网应用程序并不像传统的桌面应用程序那样完善且友好。通过 **AJAX**，因特网应用程序可以变得更完善，更友好^[20]。

第三章 用户关系可视化的分析与设计

3.1 可视化需求分析

3.1.1 需求可发展条件分析

该论文实现项目旨在帮助提高用户端获取单一人物关系数据的信息效率，提供一种有效对相关数据的分析工具，完成由数据库数据到图形数据的转换，通过其完成关系数据可视化展示的最后一环。

作为整个数据抓取展示系统的一部分，其使用数据库中已有数据进行查询展示，直面用户，拥有直观的展示方法和较为美观的界面，易用而实用

该功能以个人为主体，能够根据提供的想要等待建立联系的人的个人对相关信息并进行分析，主要包括：

(1)不适应分析：通过其以人为单位的信息收集，将信息通过此系统进行展示，进而分析出其当前是否符合此人对于某职位生活婚姻工作等状况的满足条件，经过综合分析得出不适应分析的指标。

(2)动机分析，对人文单位的机构信息进行搜集，分析出其当前的人事架构得出动机性分析的指标。

(3)可解除条件分析：对该人近期的活动地点，其所参与的项目以及他所几倍的技术进行综合的分析得出其可接触性分析的指标。

综合上述三项分析分局可发展对象的质保体系进行估计，进而综合地得出其双方简历合作的可能性，并对此进行可能的联系与商谈。

本系统应具有如下功能：

(1)查询

用户在发起查询要求后应得到相应搜索界面，可选择引力图和环形图两种表示方法，返回图表具有以用户本身为主体，和此人所有相关功能的查询数据。

(2)指定层级隐藏

对于查询出的各个节点结果，数据库中以数据表对应一相同颜色按钮，按钮通过鼠标点击方式触发，点亮或隐藏相关节点，剩余节点内容进行动态重新排列位置。其中重排方式中引力图布局按照没节点所设置引力常数的不同动态重排并达到稳定状态；环形布局方式按照剩余圆环数目进行平局后进行重新布局。

(3)数据图表缩放功能

对于有大量结果集合的图表，支持对于特定位置的缩放功能，于鼠标悬停出移动滚轮滑动，以鼠标滚轮悬停点为中心对整体图表进行缩放，使每一节点内容都能被详细展示出来

(4)节点路径展示

当鼠标悬停至某一特定节点后，位于展示图表下方位置显示此节点的上一节点以及之前节点的路径，并且光标位置附近展示该节点所相关联的详细内容。

3.1.2 用户处理逻辑

用户作为使用主题，登录模块后按所需展示环形图或引力图进行查询，引力图或圆环图展示界面，直观地对生成图形进行处理。经过综合分析得出不适应分析的指标

数据通过抓取或用户进行提交构成数据库整体数据结构。

评估者通过可发展条件分析分别进行不适应分析，动机分析和可接触分析，并针对分析结果综合得出相应结论

3.1.3 可行性分析

功能可行性：程序数据由数据库支持，本身不提供分析决策功能，分析决策经由使用者人工制定，工具本身提供高效的可视化展示方式。

实现可行性：采用较为成熟的 web 方案，动态图标库的支持使整体展示效果提高，由于 echarts 为国人团体开发维护，近期刚升级版本，技术文档等方面严重缺乏，经过大量分析后可完成所需图标要求。数据库方面使用 mysql 作为数据库后台，使用 JDBC 进行数据库与后台的链接操作，后端设计使用 java 语言在 myeclipse 进行开发，由于 mysql 轻量快捷的特点极大方便了程序所需求的多次数据库查找需求，缩短了查询时间，前端界面主要使用 javascript 技术进行开发，其已被主流浏览器所广泛支持，并在桌面端与移动端被广泛地使用，充分说明本系统在技术方面可行。

经济可行性：目前各行业招聘需求普遍采用电子简历投递，统一数据库归档方式，然而其信息提供方式普遍为通常文字条目方法，一套有效将文字人员社交信息转换为可视化图形界面的系统能大大提高信息提供效率，缩减不必要的开发成本，本系统成本较低，使用灵活，简单易学，并支持大量数据表示，在市场前景看好的经济方面可行。

操作可行性：该模块设计清晰，有良好的用户界面，操作简介，也有完善的异常处理机制和提示信息机制，对于使用人员学习成本极低，操作方面可行

法律可行性：本系统没有违反国家相关法律规定，法律方面可行。

3.2 开发方法

开发平台选用 myeclipse+mysql，web 端服务器选用 tomcat，主要逻辑通过 jsp 内置 java 标签实现，通过调用 echarts 内置 api 将数据从数据库分层取出后再分别使用 push 语句推入，完成图表的构建。

由于数据库中有数个表，所以不能一次性取出全部数据，本实现项目使用的是分次取出后处理数据再将其有所关联的下层数据为引在数据库中查找搜索再取出处理的方式完成整次查询搜索的过程。

在关系展示的设定中，通过设置引力系数，节点大小的方式，增加其易读性，并利用颜色对不同层次间的数据进行区分，使一个三层的结构能够清晰简单的体现在二维 web 网页中，同时提供了缩放浏览的方式来保证面对大量节点的图形节点信息可以被清晰地展示出来。

相比静态图形，使用纯 js 库同时也为表格展示提供了动态的表现方法，整个表格可通过选定或隐藏特定的层级按钮达到分层浏览的效果。

3.3 系统分析

整个系统前端展示为 jsp 页面，其显示效果由 echarts 以及少量 js 样式表组成，用户可通过 pc 或移动端浏览器进行访问浏览。

服务器选用 tomcat 进行部署，用户查询请求等经由 jsp 页面发送至服务器端进行查询操作。

服务器端经由 jdbc 链接 mysql 数据库进行查询读取。

以上构成了整个关系可视化系统

3.3.1 系统功能结构

功能名称：人员事件追踪

输入： 人员

功能描述： 根据用户输入的人员名称返回与该人员相关的时间追踪结果

输出：成功：返回数据到界面

失败：程序抛出异常

功能名称：机构事件追踪

输入：机构

功能描述：根据用户输入的机构名称返回与该机构相关的时间追踪结果

输出：成功：返回数据到界面

失败：程序抛出异常

功能名称：项目时间追踪

输入：项目

功能描述：根据用户输入的项目名称返回与该项目相关的时间追踪结果

输出：成功：返回数据到界面

失败：程序抛出异常

功能名称：技术事件追踪

输入： 技术

功能描述：更具用户输入的技术名称返回与该技术相关的时间追踪结果

输出：成功：返回数据到界面

失败：程序抛出异常

3.3.2 安全性分析

用户权限安全性：本系统为整体人员分析系统的一部分，整体系统安全性由后台中设置过滤器，分配权限进行保证，操作员只能在管理员分配的权限下进行操作。

系统处理事务过程按照事务处理机制执行，对不合流程的异常操作（如直接对展示界面进行访问等）进行判断，如无相关数据提交则会返回搜索界面，并短暂进行错误提示最大限度的满足数据的安全并保证系统的稳定性

3.4 系统流程图

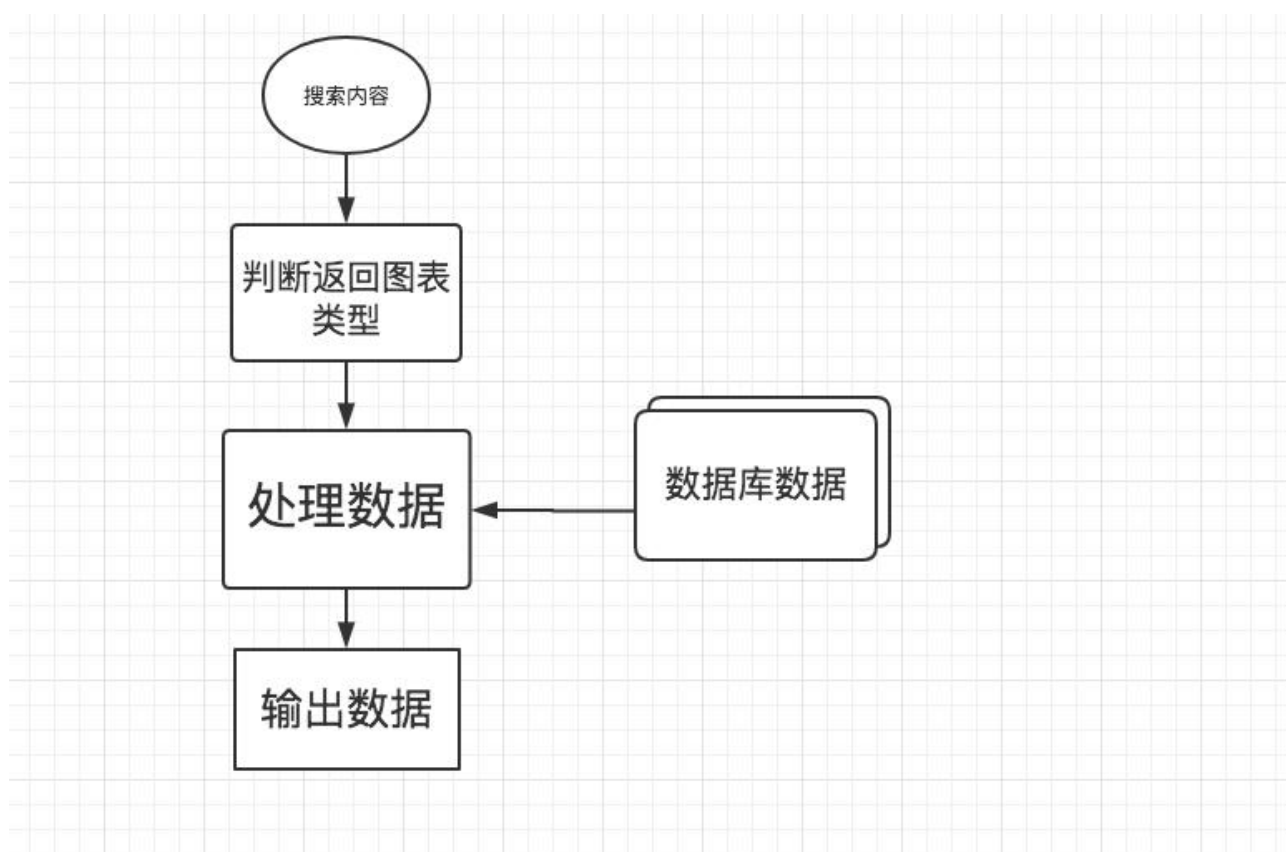


图 3.1 系统流程图

如图 3.1 系统流程图所示系统在收到搜索数据请求后会根据要求返回数据不同进行判断，系统后台收到返回数据后提取搜索关键字，通过 JDBC 与后台数据库进行搜索操作，整个搜索结束后将数据图返回至表格。

3.5 数据流程图

实际查询数据流程图

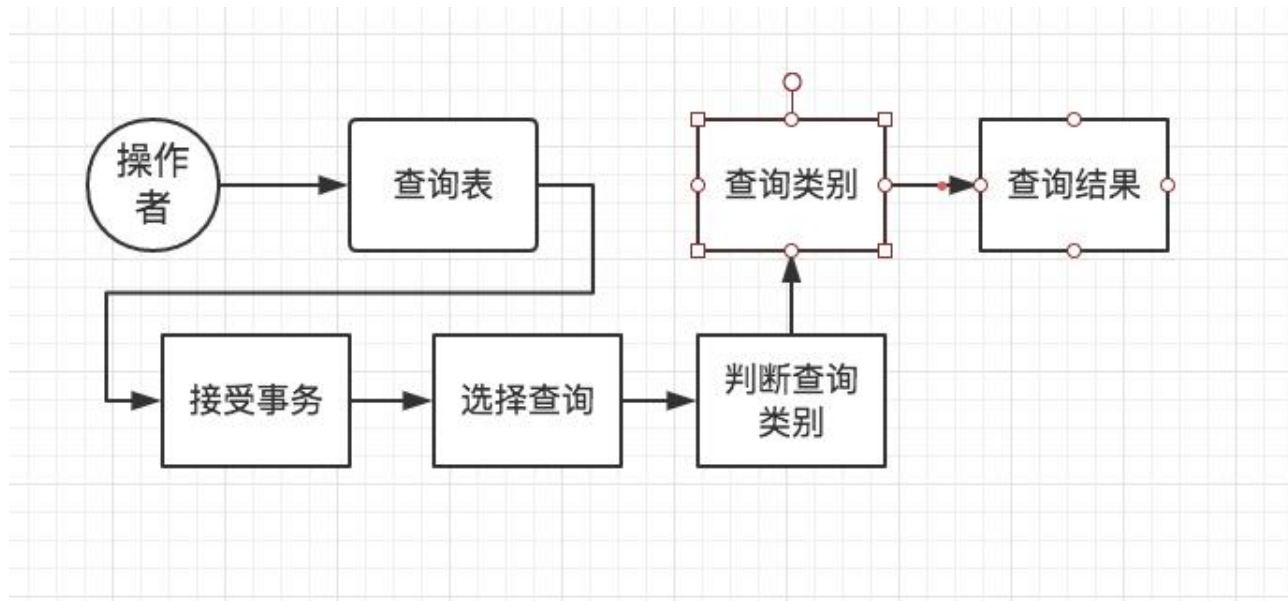


图 3.2 数据流程图

3.5 用户数据库构造

本系统数据库名为 Linkedindatabase，其中包括表，视图

使用的数据表有：人员表，教育经历表，项目经历表，工作经历表，公司经历表，具体设计如下：

人员类：

id

本族语言姓名

英文名

自我介绍

所在国家

所在地

行业

头衔

所在公司名

所在公司（机构类）的 id

毕业院校

工作经历列表 —>(为包含 工作经历类 的 List)

工作经历类:

职位（头衔）

公司名

开始时间

结束时间

公司（机构类）的 id

教育经历类:

学位

专业

学校名

开始时间

结束时间

机构类:

id

机构名

所属国家

所在地

所属行业

简介

人数规模

项目列表 （包含 项目类 的 List）

技术列表 （包含 技术类 的 List）

第四章 用户关系可视化的实现

4.1 模块功能实现介绍

最后完成后项目展示首界面如下



图 4.1 搜索界面

设计分辨实现了引力图与圆环图的两种表示方式，可通过对用户姓名进行搜索实现该用户关系的展示。

4.1.1 引力图

引力图示例如下：

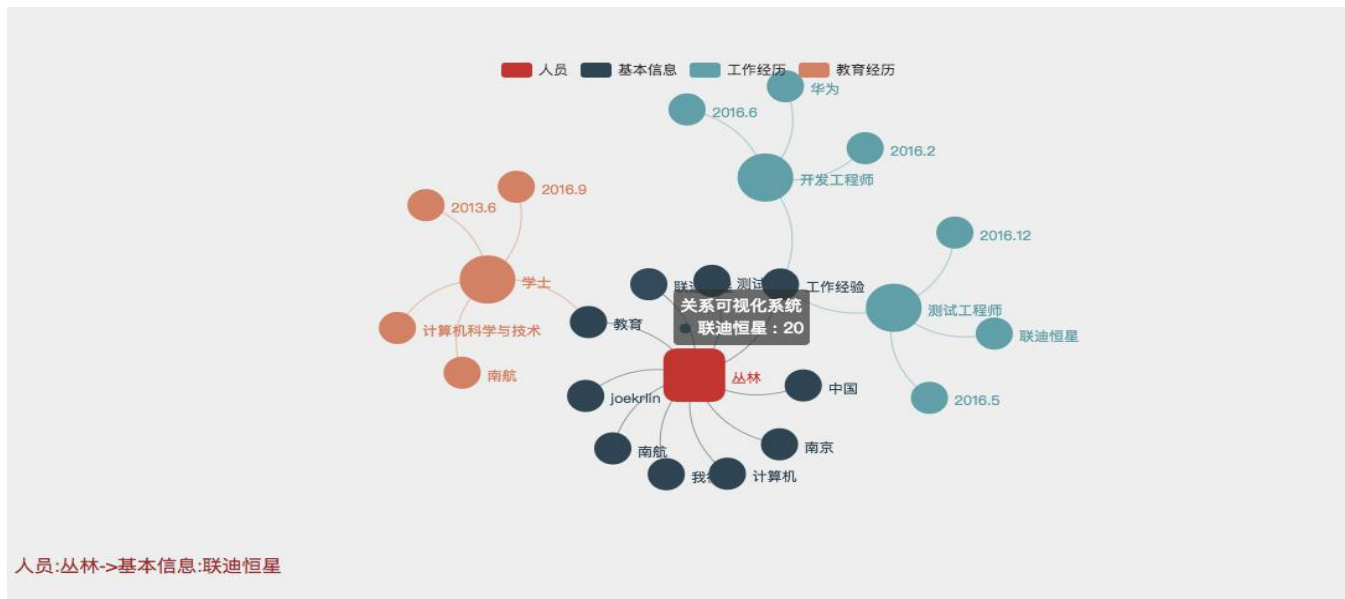


图 4.2 引力布局

引力图通过对每个引力节点进行重力相关的设置，使表格各节点在进入时由不稳定的状态通过引力与连线的模拟重力关系多次迭代验算，最终成为稳定的可视化表格，整体图形可通过鼠标滚轮缩放，鼠标悬停至节点时左下方会显示具体的关系追踪层次，针对人员情况，最多可追踪数据库中的第三季表格。通过对顶端按钮的点亮和关闭可选择隐藏或现实特定的层次级别

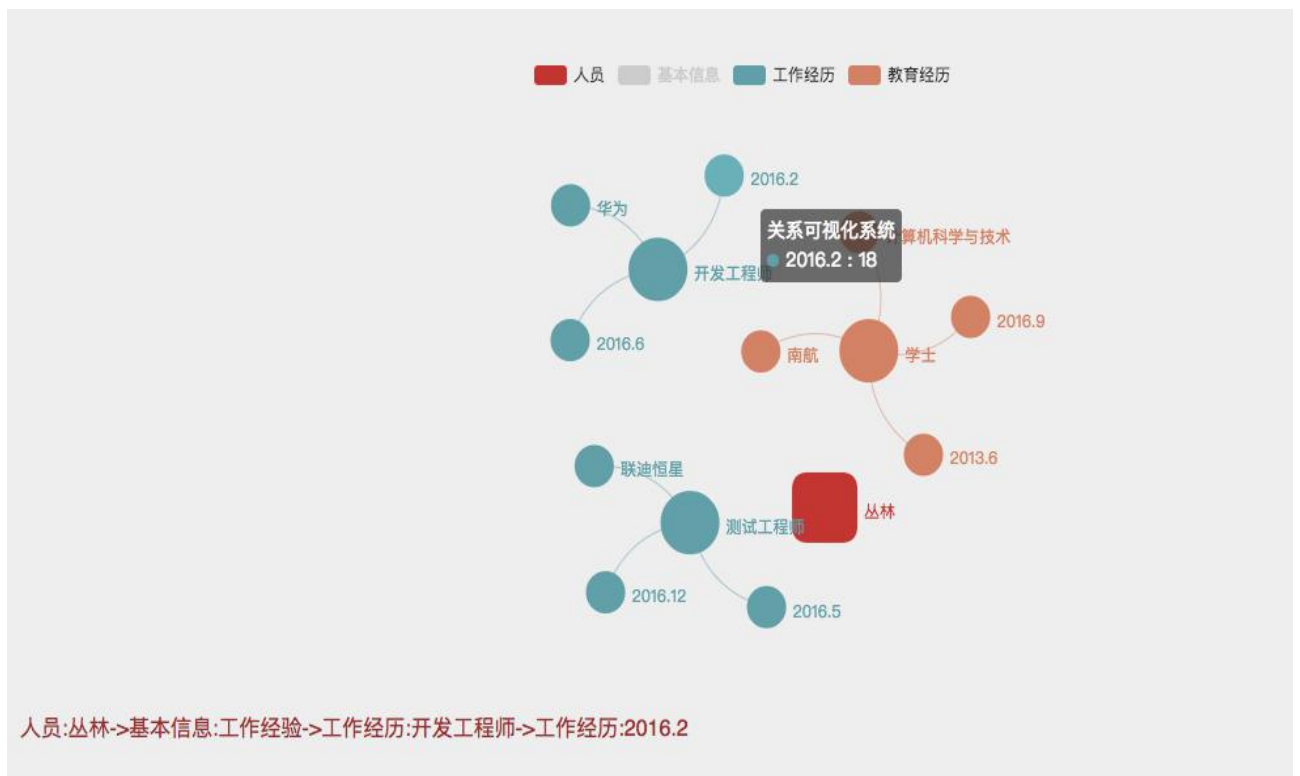


图 4.3 隐藏层级

上图为隐藏中间层而形成的仅包含后两层信息的引力图，鼠标在节点悬停后会浮出节点具体的信息展示栏

4.1.2 环形图

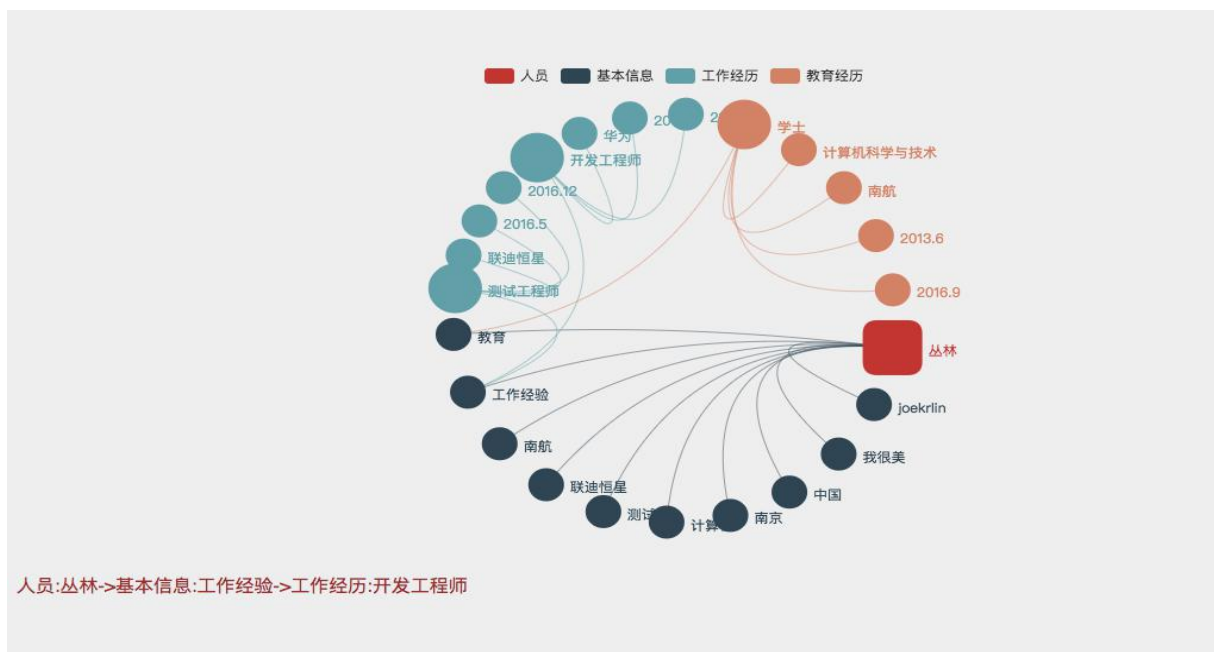


图 4.4 圆环布局

环形图对个人所有的信息数据进行了围绕排列，有关系线条体现数据间的联系，在整体数据关系的体现方面要逊色于引力布局图形，但在单独层次的关系表现效果要好于引力布局方式

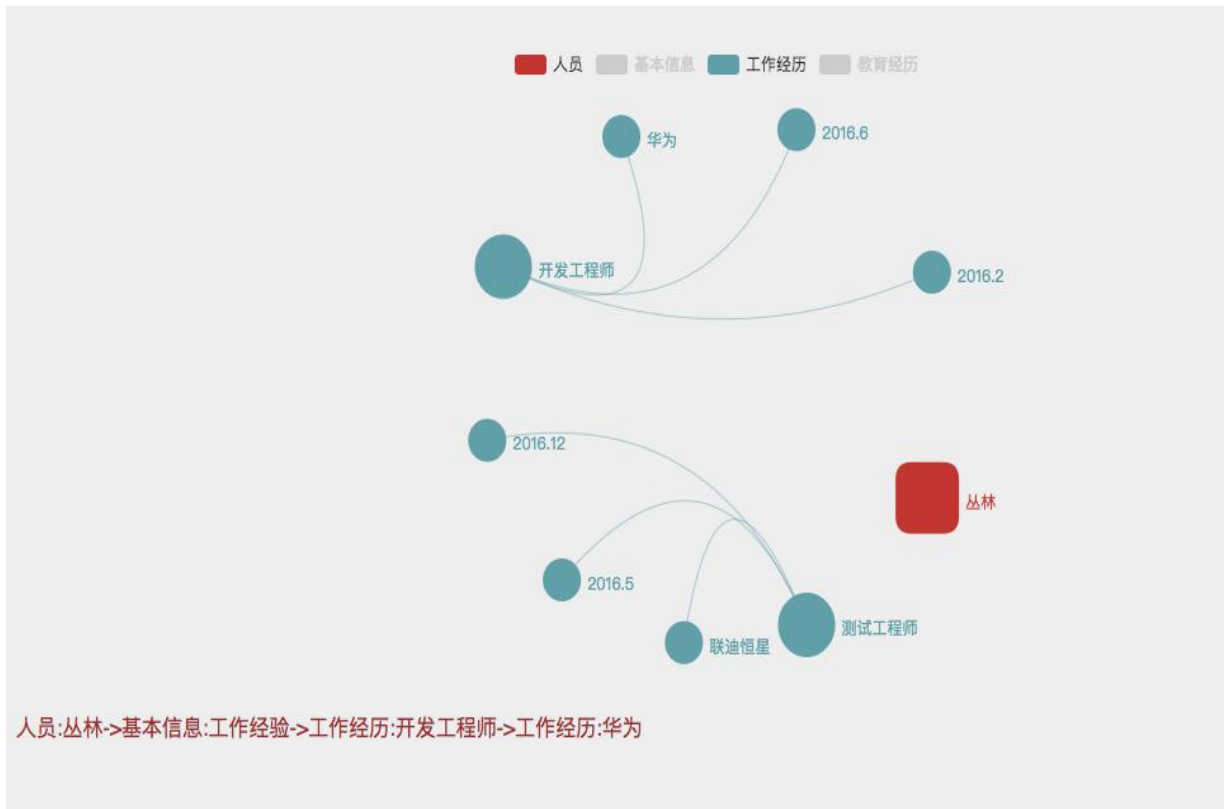


图 4.5 圆环隐藏层级

仅显示两个层级时，关系效果体现更好

4.2 细节算法与实现说明

4.2.1 后端细节设计

数据库支持：本系统所配置 MySQL 的 java 驱动版本为 5.1.39。驱动链接方式为项目文件中内置数据库链接驱动，驱动文件配置于 WebRoot/WEB-INF/lib 文件夹下。

数据库链接参数：链接参数采用标准 JDBC 链接参数配置，图 4.6 为项目具体数据库链接参数配置

```
String driverName = "com.mysql.jdbc.Driver"; //驱动名称
String DBUser = "root"; //mysql用户名
String DBPasswd = "qwer"; //mysql密码
String DBName = "Linkedindatabase"; //数据库名
```

图 4.6 数据库具体参数设置

数据库查询语句设置：本系统中 mysql 查询语句存储在 query_sql 变量中，频繁使用 select 对数据库进行查询。

数据库查询结果：数据库查询结果指针被设置保存于名为 rs 的指针中，常使用指针遍历指令对数据库返回数据进行操作。

异常处理：主体逻辑语句使用 try/catch 方法处理函数，当程勋运行出错时姜茶程勋中错误内容并返回程勋出错信息，在捕捉到错误信息后，使用 e（ArithmeticException e）对象来接收异常对象并输出异常具体信息，异常会从最近的 try catch 中输出一个后运行正常代码。

节点数据存储：节点信息被存储于名为 data 变量中（javascript 所声明变量）中，使用 data.push（echarts 内置 push 函数），对节点数据进行添加操作。本系统中所有节点分为四级，分别对应搜索个体的不同信息内容。

表 4.1 一级数据配置

数据名	变量名	数据表	数据库内名	级数
姓名	Username	User	Usrn	一
英文名	UserEnglishname	User	Usern	一
自我描述	Selfintroduction	User	Selfidr	一
国家	Country	User	Country	一
居住地	Location	User	Location	一
专业	Industry	User	Industry	一
职位	Title	User	Title	一

表 4.2 二级数据配置

数据名	变量名	数据表	数据库内名	级数
技术名称	Techname	company	Techname	二
所在公司	Company	company	Company	二
介绍	Introduction	company	Introduction	二
行业分类	Industry	company	Industry	二

表 4.3 三级数据配置

数据名	变量名	数据表	数据库内名	级数
学位	Degree	Education	Degree	三

专业	Major	Education	Major	三
学校	College	Education	College	三
开始时间	Beginntime	Education	Beginntime	三
结束时间	Overtime	Education	Overtime	三

其中单次节点内存储时数据格式为通过逗号分隔，每个设置项由关键字和设置值组成，具体设置如下：

表 4.4 节点数据配置

关键字	设置项含义	设置项配置
Id	节点编号	id:0
Category、	层级标识	Category:0
Name	节点名称	name
		'<%=rs.getString("username")%>'
Value	节点数值	Value:20
Symbol	节点显示形状	symbol : 'roundRect'
Symbolsize	节点显示大小	Symbosize:50

连线信息配置：系统中所有连线数据存储于变量 link 中，通过 link.push 添加数据表 4.5 展示了节点间连线数据配置

表 4.5 节点连线配置配置

关键字	设置项含义	设置项配置
Source	原节点 id	I+1
Target	指向节点 id	I+2

4.2.2 多级表搜索逻辑

每次搜索事件点击后，所需搜索关键字有前台 jsp 的表格设置 post 方法后后端数据通过 get 方法得到用户输入数据内容，使用所得数据在数据库中标明一级信息表中进行搜索操作，将所有搜索信息返回前端，并将数据加入节点信息，对于拥有更加详细内容的特定节点（教育信息，公司信息等），加入节点数据后使用其本身数据内容再在数据库中进行二级或三级

数据的搜索操作，设置关键信息完善后可通过改变关键字数与扩充数据库信息达到多级数据，大量对绘制节点的效果。

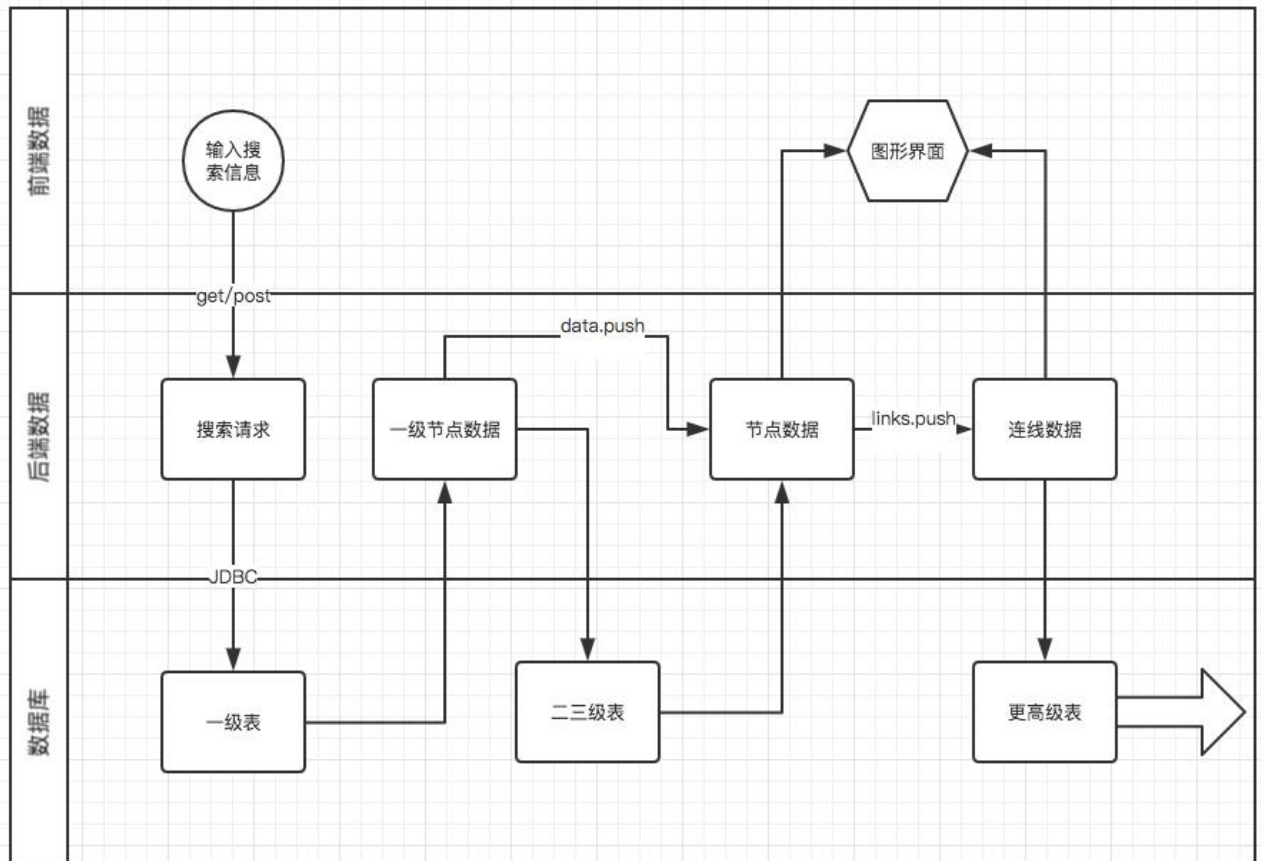


图 4.7 多级表搜索逻辑

节点数据载入完毕后，对节点间进行连线操作，连线策略为：读取所有以及表节点 id，将其与初级（用户最初节点）进行连线（使用 links.push 语句将原节点与目标节点 id 信息存入对象 links），其后对于二三级或更多级节点连线策略为读取分级二三级或更多级节点数据与本机对应的关键节点并将所有多级节点作为目标节点，其本身所对应的关键节点对应为元节点进行连线操作。图 4.8 展示了节点的连线策略，可对支持层级进行拓展从而达到图表内收录更多节点信息的效果。

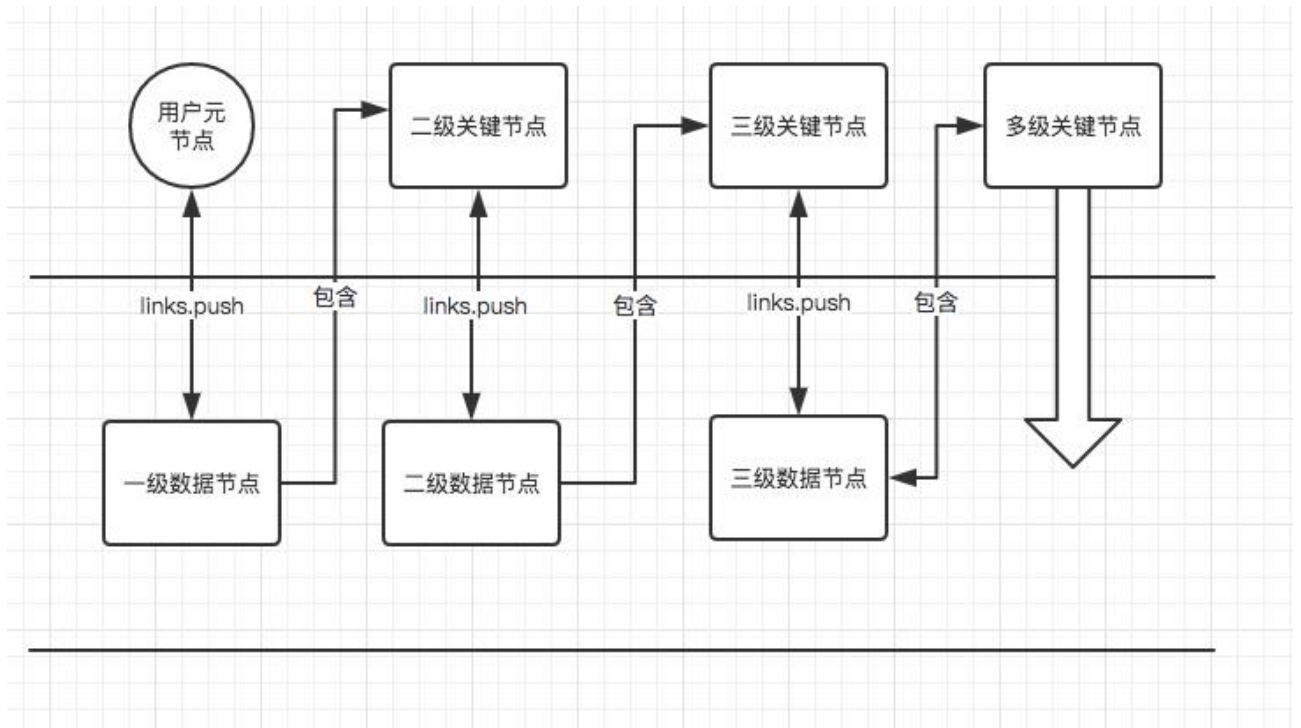


图 4.8 节点连线策略

4.2.3 前端页面细节设计

搜索界面：搜索界面使用 html 设置两个表单，分别为引力图表的设置表单与环形图表的设置表单，光标点按时使用 jquery 实现了表框发亮效果和鼠标悬停时按钮颜色加深的效果进行了浏览效果优化

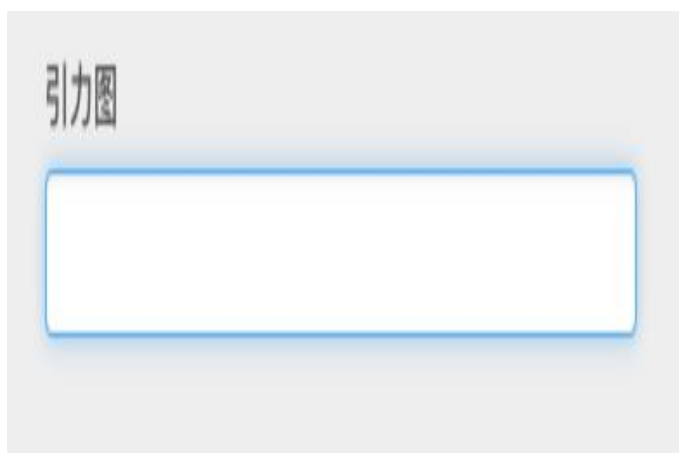


图 4.9 光标栏发亮优化

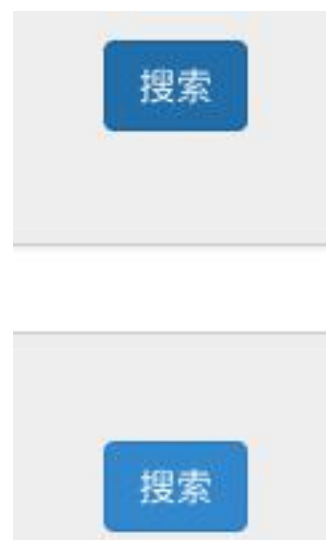


图 4.10 鼠标悬停按钮颜色加深

搜索表单通过 get/post 方法对后端进行数据传递，点按搜索按钮后会将输入栏内容发至后端进行数据搜索处理，之后会自动跳转至 findinglistall.jsp 页面并递交相关数据。

表 4.6 搜索表单配置

配置项	意义	配置值
name	表单名	Reg
meth	表单方法	Post
action	跳转路径	<%=request.getContextPath()%>/listall.jsp
class	表单种类	form-group
style	表单样式	position: fixed
border	边框属性	0

搜索表单如表 4.6 所配置，上下共设置两个，分别对应引力图表和环形图表配置搜索页面前端 js 包

表 4.7 配置所需 js 支持包

js 包名称	用途	引用方式
datatool.js	图表数据工具	<script src="js/datatool.js"></script>
echarts.js	图形库	<script src="js/echarts.js"></script>
Jquery.js	js 核心框架	<script src="js/jquery.js"></script>
Prefixfree.min.js	渐变效果包	<script src="js/prefixfree.min.js"></script>

echarts 容器的设置：在界面配置长宽固定的 dom 后通过 setoption 语句（内嵌 javascript 代码）配置 echarts 容器。配置容器后设置 echarts 支持引力图或环形图表控件，其中环形图节点大小配置使用分级依次减小方式，引力图配置方法中引力系数按照按级数级数越大引力系数越高方式，连线硬度值设为刚体对整个系统进行模拟。

表 4.8 节点显示效果配置状态

配置参数	配置值	解释
typer	Graph	关系图
name	监控管理系统	系列名称，用于 tooltip 的显示，legend 的图例筛选，在 setOption 更新数据和配置项时用于指定对应的系列。
layout	Force	于指定对应的系列。
lengedhJoverLink	True	图的布局，类型为力导图，'circular' 采用环形布局，见示例 Les Miserables
hoverAnimation	True	是否开启鼠标悬停节点的显示动画
coordinateSystem	Null	坐标系可选
xAxisIndex	0	x 轴坐标 有多种坐标系轴坐标选项
yAxisIndex	0	y 轴坐标
force	\	力引导图基本配置
nitLayout	\	力引导的初始化布局，默认使用 xy 轴的标点
repulsion	100	节点之间的斥力因子。支持数组表达斥力范围，值越大斥力越大
gravity	0.03	节点受到的向中心的引力因子。该值越大节点越往中心点靠拢。
edgeLength	80	边的两个节点之间的距离，这个距离也会受 repulsion。[10, 50]。值越小则长度越长
layoutAnimation	True	因为力引导布局会在多次迭代后才会稳定，这个参数决定是否显示布局的迭代动画，在浏览器端节点数据较多（>100）的时候不建议关闭，布局

		过程会造成浏览器假死
nodeScaleRatio	0.6	鼠标漫游缩放时节点的相应缩放比例，当设为 0 时节点不随着鼠标的缩放而缩放
Draggable	True	0 时节点不随着鼠标的缩放而缩放
focusNodeAdjacency	true	是否在鼠标移到节点上的时候突出显示节点以及节点的边和邻接节点。
symbolSize	10	也可以用数组分开表示宽和高，例如 [20, 10] 如果需要每个数据的图形大小不一样，可以设置为如下格式的回调函数：(value: Array number, params: Object) => number Array
symbol	:'roundRect'	关系图节点标记的图形。ECharts 提供的标记类型包括 'circle'(圆形), 'rect'（矩形）, 'roundRect'（圆角矩形）, 'triangle'（三角形）, 'diamond'（菱形）, 'pin'（大头针）, 'arrow'（箭头） 也可以通过 'image://url' 设置为图片，其中 url 为图片的链接。'path://' 这种方式可以任意改变颜色并且抗锯齿
edgeSymbol	'none', 'none'	边两端的标记类型，可以是一个数组分别指定两端，也可以是单个统一指定。默认不显示标记，常见的可以设置为箭头，如下：edgeSymbol: ['circle', 'arrow']
borderType	Solid	图形描边类型，默认为实线，支持 'solid'（实线）, 'dashed'(虚线), 'dotted'（点线）。

表 4.9 连线显示效果配置状态

配置参数	配置值	解释
lineStyle		关系边的公用线条样式。
type	dotted	线的类型 'solid'（实线）'dashed'（虚线）'dotted'（点线）
curveness	0.3	线条的曲线程度，从 0 到 1
opacity	1	图形透明度。支持从 0 到 1 的数字，为 0 时不绘制该图形。默认 0.5
emphasis		高亮状态
label		图形上的文本标签
show	true	是否显示标签
position	inside	标签的位置。['50%', '50%'] [x,y]
textStyle		标签的字体样式
color	#cde6c7	字体颜色
fontStyle	normal	文字字体的风格 'normal'标准 'italic'斜体 'oblique' 倾斜
fontWeight	bolder	normal'标准'bold'粗的'bolder'更粗的'lighter'更细的或 100 200 300 400..
fontFamily	sans-serif	文字的字体系列
fontSize	12	字体大小
emphasis		高亮状态
edgeLabel		线条的边缘标签
categories	name : '负载', symbol : 'rect',	symbol name: 用于和 legend 对应以及格式化 tooltip 的内容。 label 有效

第五章 总结与展望

5.1 总结

本论文主要使用 jsp 与 javascript 技术完成了对特定抓取后所得的关系型数据库系统对象数据进行读取并发掘其中其他可能存在隐形关系的分析，并以可视化的关系展示，达到了优化数据展示方式，提供更为直观地对象关系链分析工具的效果，其中编码实现思路如下：

后端系统：实现了对关系型数据路中的数据信息再构建并组织提出，设计前提供数据设计由爬虫抓取互联网数据，再由系统进行组织重新构建后将结构析出，实现了间接的数据处理与发掘工作，理论上工具支持无线层级的发掘工作，可追溯分析单个对象的各种关系链条。利用程序中数组对象二维链化关系并存储，多次，多级地访问数据信息。

前端系统：纯 javascript 满足不同端快速流畅浏览，使用纯 javascript 框架逻辑能大大减轻服务器的运行压力，将计算工作放在本地，也能大大降低对网络传输质量的要求，在计算量不变的情况下能极大增加数据信息的获取效率，分别使用引力与环形分布算法，使浏览界面信息密度合理，达到了便于阅读的效果。在面对大量节点数据的情况，能根据 pc 端与移动端的不同，分别进行相应优化工作，在不同大小的展示端口中都能获得极佳的展示效果。采用了鼠标定点缩放的方式，可针对整个图表中的任意一点进行缩放拖拽，缩放过程中所有图表采用矢量绘图方式，因而在细节处与浏览整体分布时都能达到良好的显示效果。针对特定有数值型数据的节点，采用了将节点数据隐藏，鼠标悬停后显示节点细节表格的方式。考虑到有较多层级的节点情况，在表格下方设置了展示数据关系链条的功能，使得用户主体能完整快速地获取整体信息并加以分析判断。

考虑到整体系统的可移植性，并且对数据安全性的要求较低，后端采用了直接在 jsp 页面代码中使用标签方式内嵌 java 逻辑代码，避免了和未来可能移植入系统中函数或变量间的冲突问题，直接将数据库链接代码写入 jsp 页面中，不借助 j2ee 开发框架使得整体系统轻量而便于移植，编码规则按照《阿里巴巴 java 开发规范》进行书写，注释详细，增加了整体代

码的易读性，在配置项目代码中展示全部配置接口并进行解释，使得对他人的代码交接成本大大降低，仅需进行简单修改，就能按需求配置出不同表格展示工具。

5.2 展望

本论文所实现为针对个人信息及其相关信息的可视化操作，其实现技术支持更大量数据的拓展，之后会对人员及人员间的社交关系可视化进行支持。搜索方面对特定信息的筛选功能也有很大的改进空间。系统中所实现的为引力图与环形图，针对人员情况，采用地域关系等图表分析也能达到很好效果，其主要实现方法与本论文相同，可进行配置后加入支持的功能中。

论文中所实现工具并未采用开发框架进行搭建，其稳定性与安全性均有所欠缺，搜索信息采用明文传输方式，实际使用过程中特别是网络发布使用过程中又可能被其他设备获取或监听得到相关信息，存在着一定的泄露用户隐私信息的风险。如上缺点也需要在后继版本中改进解决。

不仅仅局限于对特定结构的数据库进行分析，由主层开始，层层获取联系信息的方式也可用于其他领域，本论文所实现工具理论上可提供对于任意关系型数据库所有数据信息的抓取并将其可视化展示尾节点表格，为用户提供快捷稳定有良好易用性的可视化分析工具。

参考文献

- [1] boyd, danah; Ellison, Nicole. Social Network Sites: Definition, History, and Scholarship. Journal of Computer-Mediated Communication. 2008, 13: 210 – 230.
- [2]陈胜楠. 基于新浪微博的社交网络数据分析与可视化系统设计与实现[J]. 网络安全技术与应用, 2014(9):17-18.
- [3]李权. 面向多维数据及微博社交网络的可视分析技术的研究[D]. 清华大学, 2012.
- [4]Sun E,Rosenn I,Marlow C,et all. Gesundheit!modeling contagion through facebook news feed.Proc,ICWSM,2009,9:146-153
- [5]Cha, Meeyoung, Mislove, et al. A measurement-driven analysis of information propagation in the flickr social network[J]. Journal of Athletic Training, 2009, 45(3):215-221.
- [6]Sakaki T, Okazaki M, Matsuo Y. Earthquake shakes Twitter users:real-time event detection by social sensors[C]// International Conference on World Wide Web, WWW 2010, Raleigh, North Carolina, Usa, April. DBLP, 2010:851-860.
- [7]Ho C T, Li C T, Lin S D. Modeling and Visualizing Information Propagation in a Micro-blogging Platform[C]// International Conference on Advances in Social Networks Analysis and Mining. IEEE, 2011:328-335.
- [8]陈召群, 车向前, 谷源涛. 基于谱聚类的大规模社交网络可视化方法研究[J]. 计算机仿真, 2015, 32(9):318-322.
- [9]丁晓东. 用于网络监管的社交网络可视化分析工具的设计与实现[D]. 上海交通大学, 2007.
- [10]陈玉英. 基于机器学习的社交网络用户特征分析[D]. 北京交通大学, 2015.
- [11]赵国庆, 黄荣怀, 陆志坚. 知识可视化的理论与方法[J]. 开放教育研究,2005,11(1):23-27.
- [12]李绪, 曹磊, 付磊. 社交网络数据个性化推荐的可视化方法[J]. 计算机工程, 2014, 40(3):46-50.
- [13]李刚. 疯狂 Java 讲义（第 2 版）：电子工业出版社, 2014: 2
- [14]赵景晖. Java 程序设计：北京机械工业出版社, 2005: 1-2
- [15]Dreamweaver 动态网页文字教程 [http://soft.zol.com.cn/school/article/_4660.shtml]
- [16]Kofler M K D. Mysql[M]// MySQL. Apress, 2001:1-22.
- [17]Williams H E, Lane D. Web Database Applications with PHP and MySQL[J]. Oreilly Media, 2004.
- [18]David, Flanagan. JavaScript: The Definitive Guide (6th ed.). O'Reilly & Associates. ISBN ISBN 978-0-596-80552-4.
- [19]Nicholas, Zakas. Professional JavaScript for Web Developers (3rd ed.). Wrox. 2012. ISBN ISBN 978-1-118-02669-4.
- [20]《Ajax 基础教程》, Foundations Of Ajax 中文版, 人民邮电出版社图灵公司出版。ISBN 7-115-14481-8

致 谢

感谢夏正友老师对我的指导和帮助，这次的毕业设计代码写的十分痛快，学到了很多很多，毕业之后也会以您为榜样继续在计算机的方向学海泛舟、寻求真理，也谢谢朱学长提供的项目需求书作为参考，最终论文能顺利完成。