

北京邮电大学

大学生创新创业训练计划项目 定级检查表

项目编号： 202203005

项目类型： 创新训练☒ 创业训练☐ 创业实践☐

项目名称： 跨学科合作网络发现与推荐

项目名称（英文）： Interdisciplinary Collaboration
Network Discovery and Recommendations

项目依托学院： 计算机学院（国家示范性软件学院）

项目负责人： 胡逸同

联系电话： 13661192968

E-mail： wiederholung@outlook.com

指导教师： 王祎

E-mail： yiwang@bupt.edu.cn

起止年月： 05/2022-04/2023

填报时间： 2022 年 9 月 10 日

填写说明

- 1、 本检查表所列各项内容均须实事求是，认真填写，表达明确严谨，简明扼要。
- 2、 检查表为大 16 开本（A4），根据内容需要可自行加页，但格式须与原件一致。
- 3、 检查表填写完毕后，须在“北京邮电大学大学生创新创业计划训练平台”端口开放时间段内上传并提交，并请项目指导教师在网上进行审核。
- 4、 提交检查表前，要确定所有团队成员和指导教师已经在“北京邮电大学大学生创新创业计划训练平台”网站上注册完成，否则无法正常提交。
- 5、 检查表填写时，“项目组成员签字”“指导教师签字”“学院检查小组意见”“学校意见”无须填写。

项目名称	跨学科合作网络发现与推荐					
项目性质	创新训练 <input checked="" type="checkbox"/> 创业训练 <input type="checkbox"/> 创业实践 <input type="checkbox"/>					
项目来源	导师科研 <input checked="" type="checkbox"/> 自主探索 <input type="checkbox"/> 滚动支持 <input type="checkbox"/> 校企合作 <input type="checkbox"/> 科研院所合作 <input type="checkbox"/> 雏雁获奖 <input type="checkbox"/> 校际合作 <input type="checkbox"/>				项目起止时间	05/2022-04/2023
指导教师	王祎	职 称	教授	电话	13127958790	
项目完成人	姓名	所在学院	专业	班级	学号	联系方式
	王伊哲	国际学院	电子商务及法律	2020215111	2020213355	15353519783
	胡逸同	国际学院	电子商务及法律	2020215111	2020213350	13661192968
	巢环宇	国际学院	电子商务及法律	2020215111	2020213349	15580092683
	朱子炫	国际学院	电子商务及法律	2020215111	2020213363	13935559636
	王俊翔	国际学院	电子商务及法律	2020215111	2020213354	18910761281
项目类别	<input type="checkbox"/> 智能硬件 <input type="checkbox"/> 社交媒体 <input type="checkbox"/> 数字娱乐 <input type="checkbox"/> 通信网络 <input type="checkbox"/> 医疗健康 <input type="checkbox"/> 公共服务 <input type="checkbox"/> 电子商务 <input checked="" type="checkbox"/> 教育文化 <input type="checkbox"/> 房产家居 <input type="checkbox"/> 理论研究 <input type="checkbox"/> 机器人 <input type="checkbox"/> 无人机 <input type="checkbox"/> 智能制造 <input type="checkbox"/> 智能交通 <input type="checkbox"/> 创意设计 <input type="checkbox"/> 红旅专项 <input type="checkbox"/> 其他_____					
检索关键词	推荐系统，交叉学科，网络图谱，语义分析，网络可视化					

一、项目进展情况说明

1、项目计划要点（目标、内容、关键技术、创新点、商业模式）和调整情况

（是否按研究计划进行，若研究内容或人员有调整和变动，请说明原因）

项目前期大创小组成员每周花费七个小时左右的时间自主学习 Python、数据库、前端等等知识并自主熟悉与本项目相关的技术。本项目于 2022 年 1 月 1 日正式完成前期学习进入中期建设。

项目中期主要分为获取整合数据、数据库搭建、前端网页编写、后端服务器部署及相关功能搭建、前后端连接等等几个部分。截至 2022 年 6 月，数据的获取与整合已经完成，小组全力进行数据库的搭建与前端网页搭建的准备。本项目于 2022 年 9 月已完成基础测试版，进行功能测试。

项目后期主要工作为进一步完善项目功能排除漏洞，并进一步扩大项目涵盖学者范围。待基本功能完善之后本项目将着手准备通过算法为学者智能推荐潜在合作对象、合作领域与合作内容板块的搭建与测试。后期于 2022 年 9 月开始。

2、目前工作主要进展

本阶段项目中期任务（获取整合数据、数据库搭建、前端网页及相关功能搭建、前后端连接，**网页正式上线**等几个部分）已经完成，应用于测试阶段的网站已经可以使用，并且项目源码已经在 GitHub 上开源（Private）。

1. 数据清洗

结合教务的教师信息，小组分工，手工收集了 dblp 学者数据库内每位教师对应的 URL。

期间遇到了学者重名和英文名命名不规范的问题，小组采用人工消歧和制定英文名标准的方法解决。

2. 数据库搭建

小组讨论 MongoDB 数据库结构，提出“关系型模型”和“一对象一表的面相对象模型”两种方案。通过与指导老师探讨，考虑到今后数据的扩展和非结构化特点，采用了“面相对象的模型”，但却为日后的性能优化造成难题。

3. 数据入库

通过 dblp 提供的 API，将每位学者的个人信息和合作关系写入数据库。期间遇到了 API 数据格式变更导致数据库数据错乱，通过删库并重新适配新 API 解决。

4. 前端页面制作

本项目前端设计主要以 VUE 为载体，并使用 Echarts 表格进行辅助呈现。

本项目目前已经实现了三张表格及相关功能。首先对于中间的学者合作关系网图，此图展示了以某一学者为核心的合作关系网。用户可以通过将鼠标放置于点和线上来查询学者的基本信息与学者间的合作关系的详细信息。本项目还实现了用户点击节点切换中心查询节点的功能，用户可以点击节点来切换查询对象来查看其他学者的详细合作关系网。

对于左右两张表格，它们代表了中心节点学者的统计信息，包含中心节点学者与某一学院的合作总人数与中心节点学者与某一学院全部学者的总合作篇数。

在实现前端页面的过程中也遇到了一些问题，其大致包含：Echarts 配置文件的编写与改进，数据传递内容的统一等等。这些问题最终通过本项目组的一直努力都被成功解决。

5. 突发事件：数据库被勒索

2022.7.15，位于阿里云的 MongoDB 服务器由于未设置密码遭遇勒索攻击，要求支付 BTC 以解除加密。

转机：由于 MongoDB 采用 Docker 的部署方式，通过自动快照策略恢复了容器。

6. 业务逻辑实现

主要包含 DAO 层，并将渲染合作关系图所需数据封装成 json 文件。

由于采用“一对象一表”的数据库结构，查找学者信息需要遍历库中所有表，等待时间长，用户体验不佳。

目前通过对每张表建立索引，查找性能得到提升，但仍需探索针对多表查询的优化。

7. 合作关系图渲染

通过引入极坐标，基于合作者之间的的亲密程度，计算图中各节点的大小、相对位置和连线粗细。

8. 后端服务器部署

后端选用 Flask 最为服务器，响应 URL 并返回前端所需的 json。

期间解决了开发环境下“非本机访问”和“CORS 跨域访问”。

9. 测试调优

MongoDB 数据库，Web 服务器和后端服务器分别部署于三台不同网络的物理主机，主机间通讯均采用 HTTPS 加密协议以确保安全。

修正相关 bug 后，测试连接无误，可以正常演示功常。

至此，计划内的中期任务全部完成。与原计划相比，推迟了 1 个月。

项目进度

前期：2021.11-2022.1

学习Python、数据库、前后端等与本项目相关的技术。

中期：2022.1-2022.9

后期：2022.9-2023.4

项目后期主要工作为进一步完善项目功能排除漏洞，并进一步扩大项目涵盖学者范围。待基本功能完善之后本项目将着手准备通过算法为学者智能推荐潜在合作对象、合作领域与合作内容板块的搭建与测试。后期于2022年9月开始。

1. 数据清洗
结合教务的教师信息，小组分工，手工收集了 dblp 学者数据库内每位教师对应的 URL。期间遇到了学者重名和英文命名不规范的问题，小组采用人工消歧和制定英文命名标准的方法解决。

2. 数据库搭建
小组讨论 MongoDB 数据库结构，提出“关系型模型”和“一对多-表的面相对象模型”两种方案。通过与指导老师探讨，考虑到今后数据的扩展和非结构化特点，采用了“面相对象模型”，但却为日后的性能优化造成难题。

3. 数据入库
通过 dblp 提供的 API，将每位学者的个人信息和合作关系写入数据库。期间遇到了 API 数据格式变更导致数据库数据错乱，通过删库并重新适配新 API 解决。

4. 前端页面制作
本项目前端设计主要以 VUE 为载体，并使用 Echarts 表格进行辅助呈现。
本项目目前已经实现了三张表格及相关功能。首先对于中间的学者合作关系网图，此图展示了以某一学者为核心的合作关系网。用户可以通过将鼠标放置于点和线上来查询学者的基本信息与学者间的合作关系的详细信息。本项目还实现了用户点击节点切换中心查询节点的功能。用户可以点击节点来切换查询对象来查看其他学者的详细合作关系网。
对于左右两张表格，它们代表了中心节点学者的统计信息，包含中心节点学者与某一学院的合作总人数与中心节点学者与某一学院全部学者的总合作篇数。
在实现前端页面的过程中也遇到了一些问题，其大致包含：Echarts 配置文件的编写与改进，数据传递内容的差异等等。这些问题最终通过本项目组的一直努力都被成功解决。

5. 突发事件：数据库被勒索
2022.7.15，位于阿里云的 MongoDB 服务器由于未设置密码遭勒索攻击，要求支付 BTC 以解除加密。
转机：由于 MongoDB 采用 Docker 的部署方式，通过自动快照策略恢复了容器。

6. 业务逻辑实现
主要包含 DAO 层，并将渲染合作关系图所需数据封装成 json 文件。
由于采用“一对多-表”的数据库结构，查找学者信息需要遍历库中所有表，等待时间长，用户体验不佳。
目前通过对每张表建立索引，查找性能得到提升，但仍需探索针对多表查询的优化。

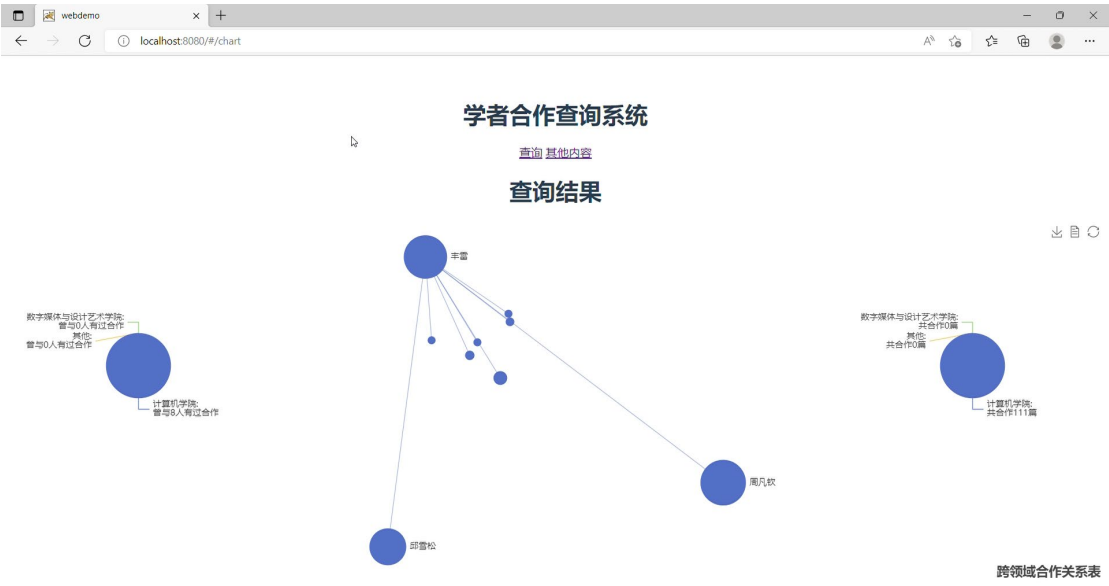
7. 合作关系图渲染
通过引入极坐标，基于合作者之间的亲疏程度，计算图中各节点的大小、相对位置和连线粗细。

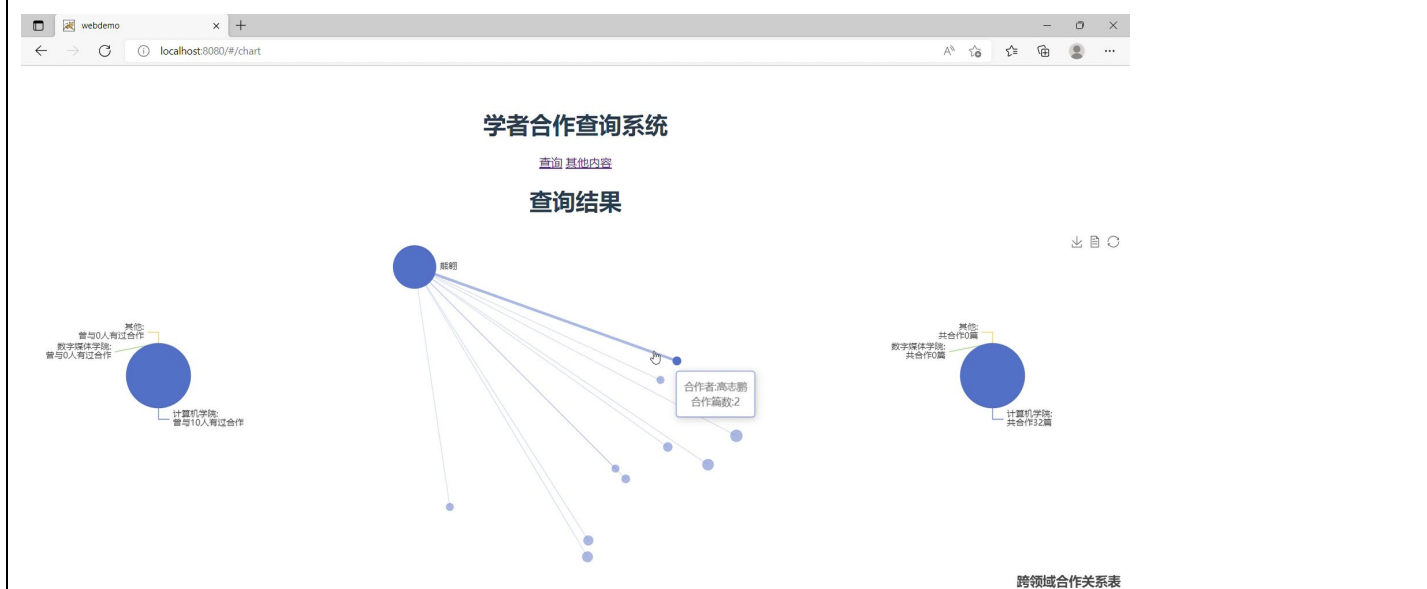
8. 后端服务器部署
后端选用 Flask 作为服务器，响应 URL 并返回前端所需的 json。
期间解决了开发环境下“非本机访问”和“CORS 跨域访问”。

9. 测试调优
MongoDB 数据库，Web 服务器和后端服务器分别部署于三台不同网络的物理主机，主机间通讯均采用 HTTPS 加密协议以确保安全。
修正相关 bug 后，测试连接无误，可以正常演示功。

至此，计划内的中期任务全部完成。与原计划相比，推迟了1个月。

3、阶段性成果（提供相关附件）





GitHub 仓库: <https://github.com/Wiederholung/Academic-Collaboration-RS>

4、目前项目经费使用情况

阿里云 OSS: 30/年

阿里云 ECS 285.18/4 年

域名 metattri.com 61/1 年

VPN 1 170/年

VPN 2 200/年

二、项目成员分工及完成情况（要求明确写出每名成员的工作内容和工作量，在项目研究过程中承担的责任和取得的成果）

姓名	完成情况	工作内容	工作量
胡逸同	已经完成	Flask 服务端编写+数据分析	每周 5h 开发
王伊哲	已经完成	算法编写+数据库调优+爬虫编写	每周 5h 开发
朱子炫	已经完成	后端+数据库	每周 5h 开发
王俊翔	已经完成	Vue+ EChart 可视化设计与编写	每周 5h 开发
巢环宇	已经完成	可视化+前端	每周 5h 开发

三、项目下一阶段工作计划及预期成果

1、 下一阶段主要研究内容和工作计划

项目下一阶段主要工作为进一步完善项目功能排除漏洞，优化数据库索引速度，并进一步扩大项目涵盖学者范围。针对每个老师的文章内容和摘要进行数据爬取，获取其相关的有用信息，以便于后续推荐系统的建立，待基本功能完善之后本项目将着

手准备通过算法为学者智能推荐潜在合作对象、合作领域与合作内容板块的搭建与测试。后期预计于 2022 年 9 月底开始。

2、 预期研究成果

主要预期成果有如下形式：

- a) 软件应用系统一套，实现系统的主要功能，在校内进行部署，在此基础上形成相应的软件著作权；
- b) 在合作分析、可视化、及推荐等关键方面，形成专利；
- c) 发表相关高水平论文 2-3 篇

3、 经费使用预算

预计版面费、专利申请及其他知识产权事务等费用：5000

书费、资料费等费用：2000

四、指导教师综合评价

本项目进展顺利，在项目开展的过程中，项目组同学能够积极学习新的技术知识体系，提出解决问题的创新思路，在项目执行的过程中，展示出较强的执行力，针对项目中出现的状况，能够迅速响应和解决，技术能力和创新能力都得到了极大的提升。目前的系统原型较好的实现了预期目标，能够保证预期成果的完美实现，建议继续努力，实现项目的完美交付和项目的结题。

项目组成员 签字	胡逸同 王伊越 王俊翔 朱子炫 巢环宇	2022 年 09 月 13 日
指导教师签字	王祎 张高唯	2022 年 09 月 13 日
五、评审意见		
专家组评价意见：		
基地主任签字： 年 月 日		
学校审批意见：		
负责人签字： 盖章 年 月 日		