

基于协同过滤算法的高校社团推荐系统的设计与实现

骆 伟, 殷宏涛, 陶 琛

(大连东软信息学院软件工程系, 辽宁 大连 116023)

✉luowei@neusoft.edu.cn; 208105166@qq.com; 709510257@qq.com



摘 要: 高校中的社团较多, 如何能让大学生快速找到感兴趣的社团, 成为社团管理者面临的主要问题。为了向大学生推荐可能喜欢的社团, 本项目通过基于协同过滤的推荐算法, 找到共同出现的频率来计算语义的相似度, 并通过计算空间向量的夹角余弦值进而计算文本之间的相似度, 能够在提高社团管理者工作效率的同时, 为需要加入社团的学生和对社团文化有浓厚兴趣的学生提供更加全面的信息。实验结果表明, 当推荐项目数量为10时, 该方法的召回率、准确率和F1值分别提高了12.81%、7.65%和14.51%, 表明基于协同过滤的推荐算法可有效提高推荐结果。

关键词: 社团管理; 推荐算法; 语义相似度; 余弦相似度

中图分类号: TP311.5 **文献标识码:** A

Design and Implementation of College Clubs Recommendation System based on Collaborative Filtering Algorithm

LUO Wei, YIN Hongtao, TAO Chen

(Department of Software Engineering, Dalian Neusoft University of Information, Dalian 116023, China)

✉luowei@neusoft.edu.cn; 208105166@qq.com; 709510257@qq.com

Abstract: There are many clubs in colleges and universities. How to make college students quickly find clubs of their interests has become the main problem faced by club managers. In order to recommend clubs that college students might like, this paper proposes to use a collaborative filtering recommendation algorithm to find the frequency of common occurrences for calculating semantic similarity. The similarity between texts is calculated by calculating the cosine of the space vector angle. It can not only improve the work efficiency of club managers, but also provide more comprehensive information for students who need to join the club and students who are very interested in club culture. Experimental results show that when the number of recommended items is 10, the recall rate, the accuracy rate and F1 value of this method are increased by 12.81%, 7.65% and 14.51% respectively, indicating that the recommendation algorithm based on collaborative filtering effectively improves recommendation results.

Keywords: club management; recommendation algorithm; semantic similarity; cosine similarity

1 引言(Introduction)

高校社团管理系统的开发目标是为大学生、社团管理层、社团负责教师提供一个社团管理的平台。本系统通过对语义相似度和余弦相似度的计算, 基于物品的协同过滤算法, 完成对用户的个性化推荐^[1]。本系统能够为广大师生用户

群提供一个学校社团信息分享的途径和一种新的社团管理模式, 让社团信息更加集中化, 为想要加入社团的学生提供更加全面准确的社团信息以及便捷的社团信息申请方式; 为社团管理者提供一个统一、全面、便捷的社团管理途径, 让社团招新、社团活动申请等从纸质化办公环境和繁杂的数据统

计中摆脱出来。本系统根据社团需求进行系统性的分析，使社团信息更加全面化，让社团事务的管理更加便捷。

2 系统架构及技术(System architecture)

高校社团管理系统采用B/S网络结构模式，以及前后端分离的开发形式。前端采用当下流行的Vue框架^[1]。Vue是一种前端渐进式的轻量级框架，采用MVVM(Model-View-ViewModel)软件架构的设计模式，实现DOM监听和数据的双向绑定，真正实现了事件驱动编程。后端使用的是Spring Boot框架^[2]。Spring Boot采用面向切面编程和声明式的编程模式，集成优化了Spring框架，只需要添加配置文件，就能把相应的组件加载进来。它以注解的形式完成了配置文件的添加，实现了自动化配置，提高了开发的效率。系统架构图如图1所示。

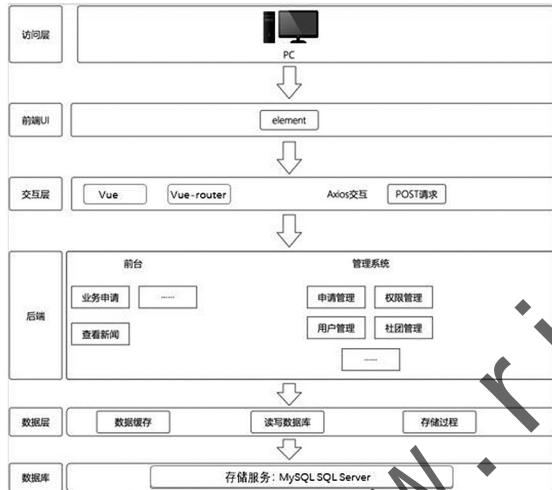


图1 系统架构图

Fig.1 System architecture diagram

3 系统需求设计(System requirement design)

3.1 系统功能结构图

根据需求调研结果确定本系统主要包括三个功能模块，分别为超级管理员、普通管理员和普通用户。系统功能结构图如图2所示。

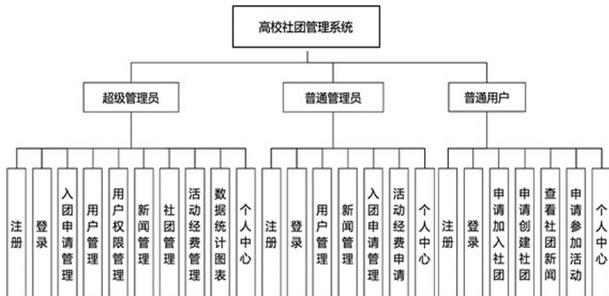


图2 系统功能结构图

Fig.2 System function structure diagram

3.2 数据库设计

概念模型(E-R图)如图3所示，共包括六个实体，分别是用户、活动与经费、社团、新闻、用户社团申请、用户活动申请。E-R图表示实体之间的类型、属性和联系方式。

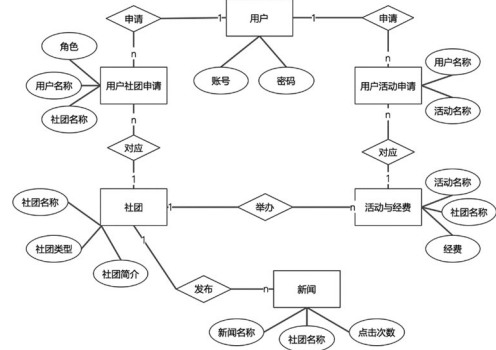


图3 E-R图

Fig.3 E-R diagram

通过对领域对象、业务对象及对象间的关联进行分析，设计了本系统的数据库用户信息表、新闻信息表、社团信息表、活动与经费信息表、用户与社团关系信息表、用户与活动关系信息表。

3.3 系统用例概述

高层用例应满足用户基本业务需求的用例概述。这些用例从教师(超级管理员)操作、社长(管理员)操作、普通用户操作三个方面完成。高层用例图如图4所示。

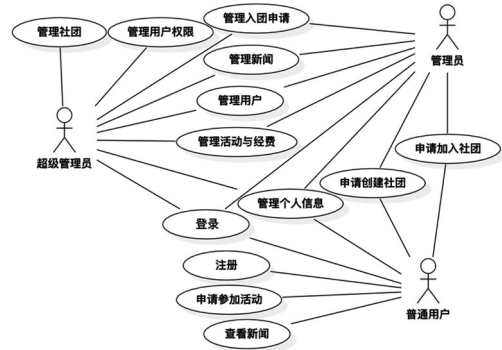


图4 高层用例图

Fig.4 High-level use case diagram

4 设计与实现(Design and implementation)

因本文的篇幅限制，不能一一介绍本系统各个功能代码和运行情况截图，且一些功能的大体执行思路相似或相对常见，所以只对系统的核心功能进行介绍和展示。

4.1 创建社团

用户可以点击创建社团，进入创建社团页面，填写社团信息创建社团。在创建社团过程中，添加一条社团信息的同时，需要向用户社团关系表中添加当前用户的用户社团关系

信息。调用Axios封装的addGroup接口，请求参数为param，向后台提交请求后接收返回数据。创建社团如图5所示。



图5 创建社团图

Fig.5 Diagram of creating clubs

4.2 社团管理

社团管理如图6所示。用户可以点击社团管理，进入社团管理页面，审批创建社团的申请。



图6 社团管理图

Fig.6 Diagram of club management

4.3 关键技术难点和解决方案

(1)JavaScript关键字Let的使用问题

在使用Let声明变量时发现声明的变量无效，导致后台无法接收到数据。出现的问题如图7所示。

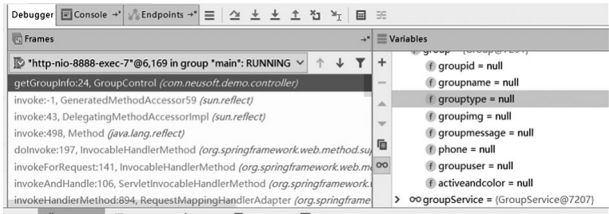


图7 出现的问题I

Fig.7 Problem I

主要问题为ES6可以使用Let关键字来实现块级作用域。因为Let声明变量后变量只在Let所在的代码块内有效，所以If语句中Let声明的param变量只在当前If语句中生效，后台无法

接收到请求参数。

(2)Vue组件中Style的Scoped属性

Scoped属性的作用是使得当前Vue组件中的样式不会影响到其他Vue组件的样式。在使用Scoped属性时，部分Element-UI样式会失效，但是为了不“污染”其他页面，不得不使用Scoped属性，这时需要自己去手动重写组件的样式。但是我们会发现重写样式时CSS代码无效，这时可以使用Vue深度选择器、/deep/深度选择器。若不写/deep/，设置的宽高是不会生效的。出现的问题如图8所示。

```
/deep/ .el-upload-dragger {
width: 250px;
height: 150px;
}
```

图8 出现的问题II

Fig.8 Problem II

(3)Axios的跨域请求问题

浏览器的安全限制限制了当前地址向非同源地址发送AJAX请求，即限制了跨域访问。本项目浏览器端的地址为http://localhost:8080，服务器端的地址为http://127.0.0.1:8888，因为浏览器同源策略的限制，即协议、域名、端口全部相同，所以浏览器端请求服务器端属于跨域访问。为解决跨域访问问题，我们在后端添加了CorsConfig过滤器配置类，进而允许跨域请求。

(4)推荐算法^[3]

基于统计经验的分布，假设认为具有相似词汇的句子应该拥有相似的语义。通过统计词汇在文本中共同出现的频率来计算语义的相似度，以计算机可读的向量形式来表达文本是该假设的主要思路。

在本系统中主要采用的是余弦相似度算法^[4]，采用构建空间向量模型的方式，通过计算空间向量的夹角余弦值进而计算文本之间的相似度^[5]。在系统中将两条新闻的标题进行分词，将两个句子的分词结果进行去重合并，将去重合并结果($T_1, T_2, \dots, T_n$)看作n维坐标系中的坐标轴，分别统计两个句子对应去重合并结果的词频($W_1, W_2, \dots, W_n$)作为坐标值，列出两个句子的词频向量，利用两个句子的向量夹角余弦值来确定两个句子的相似度，从而进行推荐^[6]。向量夹角余弦值越接近1，说明向量夹角越接近0，两个句子的语义就越相似。计算方法如下：

$$\cos \theta = \frac{\sum_{i=1}^n (x_i \times y_i)}{\sqrt{\sum_{i=1}^n (x_i)^2} \times \sqrt{\sum_{i=1}^n (y_i)^2}}$$

用户登录后, 高校社团管理系统首先收集用户的注册信息、浏览记录信息等数据, 运用基于物品的过滤协同算法, 使用余弦相似度算法计算出用户可能感兴趣的社团, 最后响应到前端个性化推荐模块。推荐流程如图9所示。

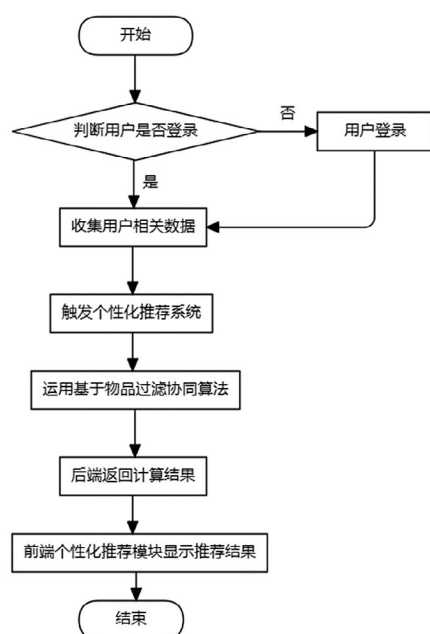


图9 推荐流程图

Fig.9 Flow chart of recommendation

5 结论(Conclusion)

社团生活是校园生活中不可或缺的一部分, 本系统将线下的部分社团事务转移到了线上, 在丰富社团生活的同时为

社团相关事务的申请和管理提供便捷。在未来的开发中我们将继续完善和改进推荐算法, 引入基于情感的相似度计算^[7]和基于信任的推荐算法^[8], 从而提高推荐的准确度。

参考文献(References)

- [1] PARK D H, KIM H K, CHOI I Y, et al. A literature review and classification of recommender systems research[J]. Expert Systems with Applications, 2012, 39(11):10059–10072.
- [2] 冯传波, 彭章友, 张钟浩. 基于Vue.js的移动应用可视化平台的研究[J]. 工业控制计算机, 2019, 32(05):106–107.
- [3] 刘敏. 基于SpringBoot框架社交网络平台的设计与实现[D]. 长沙: 湖南大学, 2018.
- [4] 刘建国, 周涛, 汪秉宏. 个性化推荐系统的研究进展[J]. 自然科学进展, 2009, 019(001):1–15.
- [5] 李连, 朱爱红, 苏涛. 一种改进的基于向量空间文本相似度算法的研究与实现[J]. 计算机应用与软件, 2012(02):282–284.
- [6] 张根宇. 基于TF-IDF和余弦相似度的文本相似度算法研究和优化策略[J]. 中国科技成果, 2019(16):25–26.
- [7] 余燕萍. 基于相似用户情感分析的推荐算法研究[J]. 科技广场, 2016(8):13–16.
- [8] 陈婷, 朱青, 周梦溪, 等. 社交网络环境下基于信任的推荐算法[J]. 软件学报, 2017, 028(003):721–731.

作者简介:

骆伟(1980–), 男, 硕士, 副教授. 研究领域: 人工智能.
殷宏涛(2000–), 男, 本科生. 研究领域: 软件工程.
陶琛(1999–), 男, 本科生. 研究领域: 软件工程.

(上接第58页)

- [2] 姚军, 党智军, 胡琳卿. VR技术在高校校园中的应用[J]. 中国科技信息, 2019(10):69–70.
- [3] w3sf. 走进3D的奇妙世界[EB/OL]. (2019–12–05) [2021–11–27]. <https://zhuanlan.zhihu.com/p/95463367>.
- [4] 王延朝. 基于Krpno的三维全景系统的开发和应用[D]. 上海: 华东师范大学, 2012.
- [5] (美) 乔斯·德克森. Three.js开发指南(基于WebGL和HTML5在网页上渲染3D图形和动画)[M]. 3版. 周翀, 张薇, 译. 北京: 机械工业出版社, 2019:24.
- [6] 赵海鹏, 周杨, 卞和方. 基于Three.js的三维虚拟校园系统设计与实现[J]. 兰州交通大学学报, 2019, 38(3):85–94.
- [7] KRAIS24. OpenGL相机[EB/OL]. (2021–07–29) [2021–11–24].

https://blog.csdn.net/krais_wk/article/details/119182495.

- [8] LIU D W, PENG J, WANG Y Y, et al. Implementation of interactive three-dimensional visualization of air pollutants using WebGL[J]. Environmental Modelling & Software, 2019, 114(04):188–194.
- [9] 如何用Three.js快速实现全景图[EB/OL]. (2019–11–08) [2021–11–24]. <https://zhuanlan.zhihu.com/p/40881782>.

作者简介:

刘晓林(1978–), 男, 硕士, 高级工程师. 研究领域: 移动应用开发, 虚拟现实, IT项目管理.
牛德雄(1965–), 男, 硕士, 副教授. 研究领域: Web开发, 软件工程, 需求工程.