

# Talent Information Flow Analysis System Based on Big Data

Zhi Sui, Jie Cao, Liwen Wang

Institute of Information Engineering, Xuzhou University of Technology, Xuzhou Jiangsu  
Email: 1006870014@qq.com

Received: Apr. 29<sup>th</sup>, 2020; accepted: May 14<sup>th</sup>, 2020; published: May 25<sup>th</sup>, 2020

## Abstract

The talent information flow analysis system based on big data systematically collects the basic talent information of the region. After sorting, the economic characteristics of the floating population include variables such as employment status, occupational attributes, unit nature, employment income and housing expenditure [1]. The main factors of these influences are extracted, and the common attributes are extracted. They can be classified according to the regional characteristics, economic impact, time, gender and other factors during the flow of talents. They can be simply divided into stable talents, mobile talents, and regression talents as the flow of talent information. The training sample set based on k-nearest neighbor algorithm analyzes the newly imported talent information, divides the talents into different categories, provides data support for decision makers, and the talent information flow analysis system based on big data provides decision makers with a visual, convenient and efficient analysis system. The system design uses UML use case diagrams for software requirements analysis, and the software system design uses package diagrams for software system design.

## Keywords

Talent Information Flow, k-Nearest Neighbor Algorithm, Data Support, UML

# 基于大数据的人才信息流向分析系统

睢 治, 曹 杰, 王立文

徐州工程学院信息工程学院, 江苏 徐州  
Email: 1006870014@qq.com

收稿日期: 2020年4月29日; 录用日期: 2020年5月14日; 发布日期: 2020年5月25日

## 摘 要

基于大数据的人才信息流向分析系统对地区的基本人才信息进行系统的收集、整理后, 流动人口的经济特征包括就业身份、职业属性、单位性质、就业收入与住房支出等变量[1], 将这些影响的主要因素提取,

提取共有属性,可以根据人才流动时的地区特点、经济影响、时间、性别等因素进行分类,简单划分成稳定型人才、流动型人才、回归型人才等作为人才信息流动的训练样本集,基于k-近邻算法对新导入的人才信息进行分析,将人才划分到不同种类中,为决策者提供数据支持。基于大数据的人才信息流向分析系统为决策者们提供一个可视化方便、效率高的分析系统,系统设计使用UML用例图进行软件需求分析,使用包图进行软件体系设计。

## 关键词

人才信息流向, k-近邻算法, 数据支持, UML

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

随着经济的发展还有城市化的发展,各地的人才流动也变得更加剧烈,各地的人才都在选择自己的择业方向,地点。人才大量的流动无论对企业,还是对当地的政府都是一项巨大的考验,针对本年的高校毕业生,各个城市之间的流动人才,这些不同专业领域的人才在本地的稳定性成为了令人困扰的难题,特别是对企业、单位的决策人来说,如果能够及时地了解到当地稳定的人才信息、数量,就可以根据这些信息做出更有利发展的决策。当然对于企业招聘过程中选择人才也有一定参考价值,基于这些原因一款基于大数据的人才信息流向分析系统必不可少。

基于大数据的人才信息流向分析系统依据人才流动科学中对人才流动的主要因素分析,建立人才流动信息训练集,然后利用机器学习中的k-近邻算法来进行新的人才的划分,将每年新的人才划分到稳定型人才、流动型人才、回归型人才中,同时也从人才流动信息入手,通过查看不同人才的分布,流动方式,用可视化的方式展示给用户等方式为用户进行人才招聘提供更多的决策支持。用户们可以看到以散点图动态展示不同领域人才的区域分布,以柱状图展示的该地区人才流入情况,以扇形图展示的不同学历的人才分布,管理员则可以进行用户管理,管理用户信息。

基于大数据的人才信息流向分析系统对解决各地企业和政府对本地的具体发展方向具有很高的数据支持作用,可以辅助他们更好地进行决策。本系统使用UML作为建模工具[2]。

## 2. 课题需求

- 1) 用户注册后完成后,即可以正常使用系统功能。
- 2) 用户可以查看人才在不同地域的分布,以散点图动态展示。
- 3) 用户可以查看地区人才的流动情况,将以柱状图展示。
- 4) 用户可以查看地区中不同学历的人才分布,将以扇形图展示。
- 5) 用户可以检索领域,查看不同专业领域人才流动情况。
- 6) 用户可以检索年份,查看不同年份人才学历分布情况。
- 7) 系统管理员进行用户管理。

## 3. 需求分析建模

基于大数据的人才信息流向分析系统具体业务围绕用户展开。管理员是系统的管理者。

可以使用业务用例图对整个系统进行业务分析建模[3]，见图 1。

管理员负责系统的用户管理，权限管理。用户注册完成后可以正常进行系统功能使用。用户可以根据自己的需求查看不同的可视化信息，包括人才分布信息，人才流动信息等。

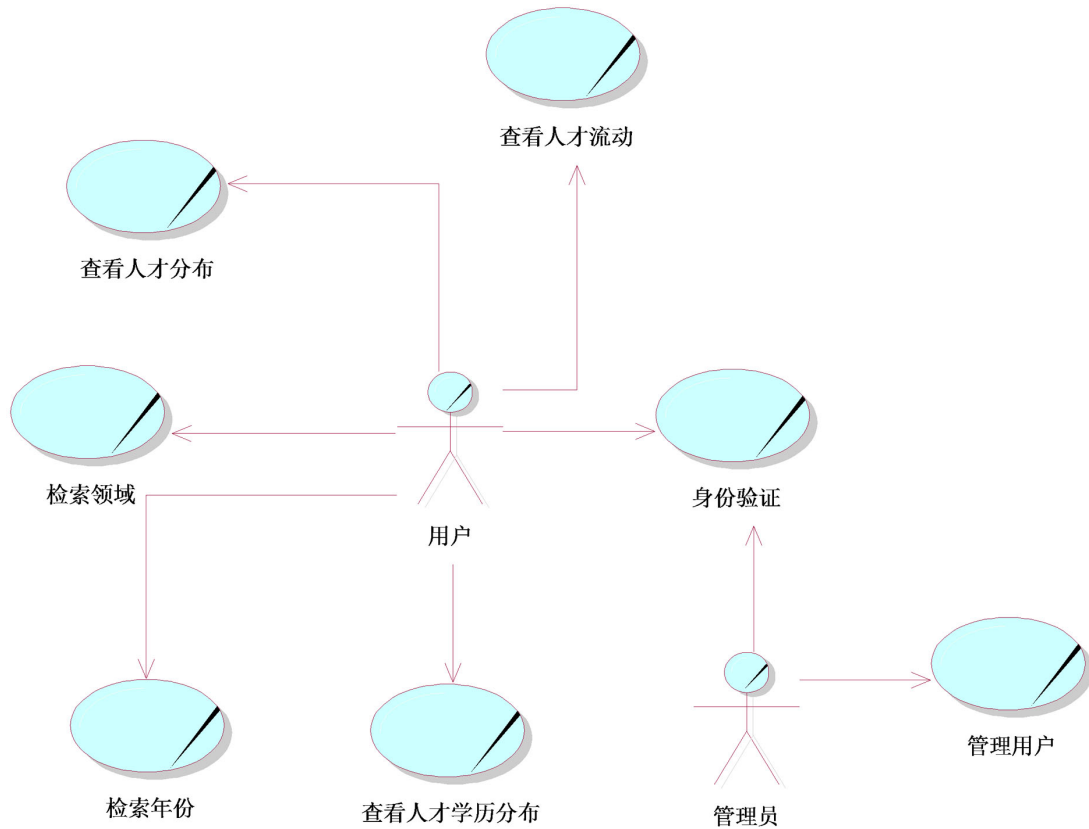


Figure 1. System business use case diagram  
图 1. 系统业务用例图

#### 4. 系统分析建模

在对人才流动因素进行系统的学习后，依据人才的地区特点，经济影响，时间，性别，学习经历等基本信息提取，作为进行样本集的分类，分别对稳定型人才，流动型人才，回归型人才种类进行权重的赋值，每个数据集都有自己的中心节点，如图 2 所示，不同的点可以很清晰的展示在坐标轴上，紫色的三角是每个分类的中点，当新的人才信息输入时，将会计算距离不同种类的中心距离，然后将新的人才归类到具体的人才种类中。

使用 KNN 近邻填补算法，获得了更好的结果，填补值更接近原始值，这样得到的数值也更贴近于实际获得的数据值[5]。

用户信息：存储用户的数据信息，主要属性包括姓名、联系方式等信息。

人才分布信息：存储人才分布的数据信息，主要属性包括专业领域、人数、地点等信息。

人才流动信息：存储人才流动的数据信息，主要属性包括专业领域、年份、流动方式、人数等信息。

人才学历分布信息：存储人才学历分布的数据信息，包括学历程度，年份，人数等信息。

账户信息：存储账户的数据信息，主要属性包括账号、密码等信息。

管理员信息：用于存储平台管理员相关的数据信息，主要属性包括工号、姓名等信息。

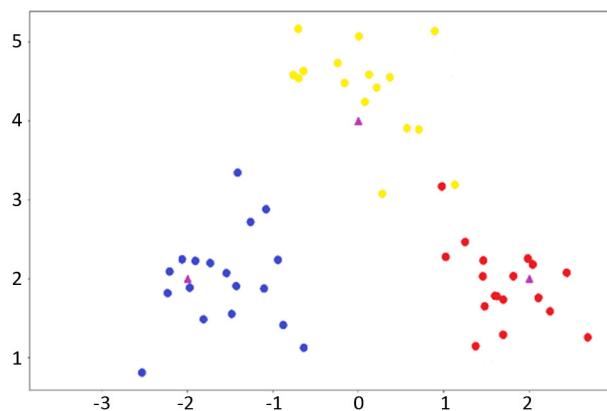


Figure 2. Talent classification k-nearest neighbor algorithm classification diagram

图 2. 人才分类 k-近邻算法分类图

使用分析类图对整个系统进行系统分析建模，见图 3。

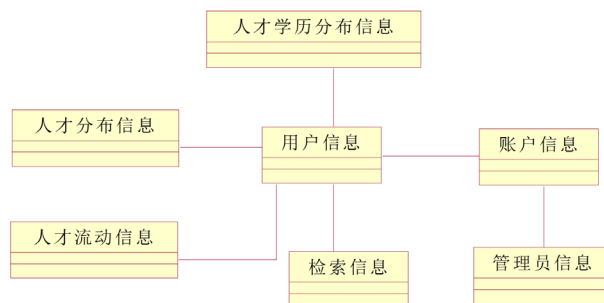


Figure 3. System analysis class diagram

图 3. 系统分析类图

基于大数据的人才信息流向分析系统设计体系图如图 4 所示。

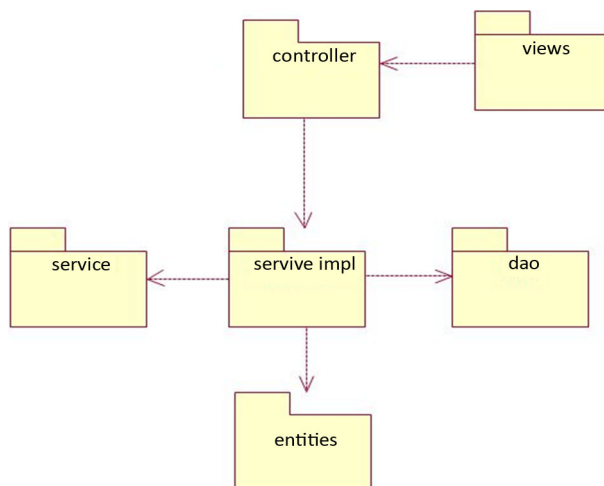


Figure 4. Architecture design package diagram

图 4. 体系结构设计包图

基于大数据的人才信息流向分析系统功能架构图如图 5 所示。

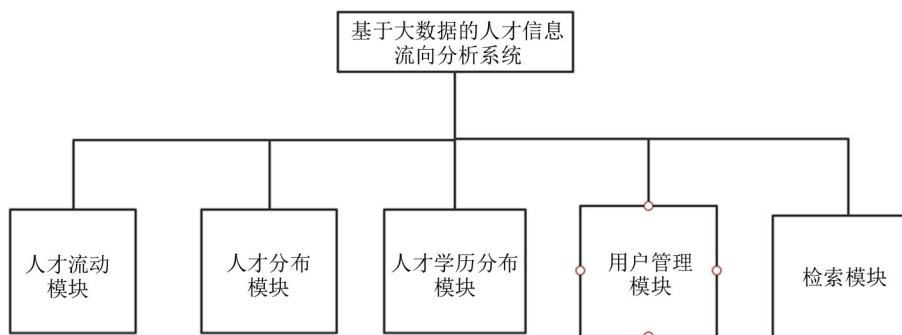


Figure 5. Software functional architecture design diagram  
图 5. 软件功能架构设计图

### 5. 功能设计

查看人才流动信息模型需要 Trans、User 类共同实现。其中 Trans 类为人才流动信息类，包含人才流动信息相关的信息，User 类为用户信息。

Trans Dao 实现了查看人才流动信息功能中的数据访问，实现人才流动信息查询 Trans Query 方法。

Trans Service 实现了具体的查看人才流动信息业务逻辑功能，主要是查询人才流动信息 Get Trans。

Trans Views 是 Web 端的页面，Web 应用开发中的网站后台管理主要包括操作管理、权限管理、实体管理以及反馈管理等四个活动内容[4]。

### 6. 软件功能测试

用户点击流动信息界面，可以查看到默认领域的人才流动情况，以柱状图进行展示，用户在查询领域输入框可以输入想要查看的人才领域，点击查询，可以查询到输入领域的人才流动信息，查看人才流动功能界面见图 6。

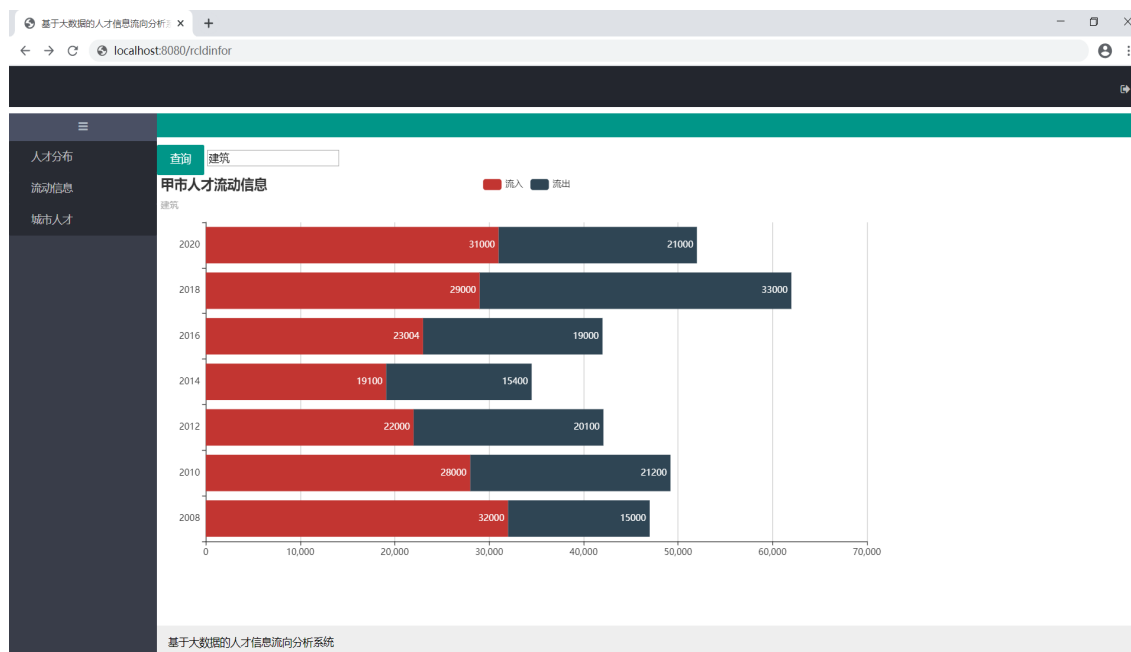


Figure 6. View the interface diagram of talent flow function  
图 6. 查看人才流动功能界面图

用户登录后可以点击城市人才，进入城市人才学历分布界面，可以看到该城市人才学历分布信息，以饼状图的形式展示，查看城市人才分布界面展示如图 7 所示。

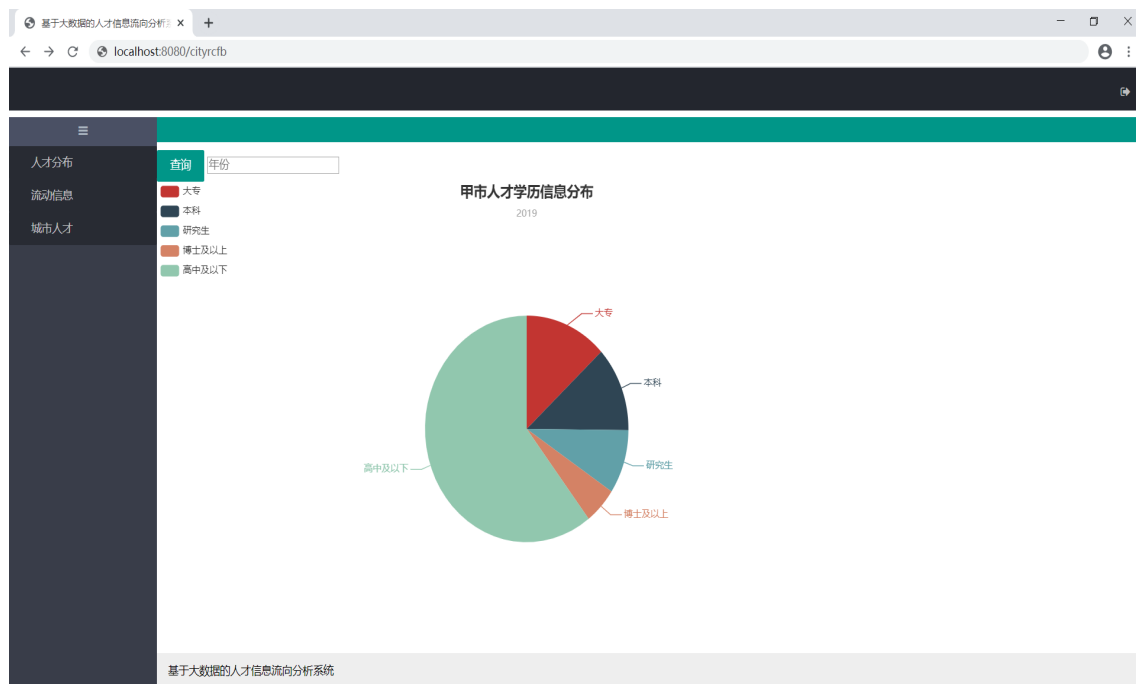


Figure 7. View the city talent distribution interface map  
图 7. 查看城市人才分布界面图

管理员登录后可进入权限界面查看使用本系统的所有用户界面操作权限信息，并且可以对每个用户权限进行分配，权限管理界面展示如图 8 所示。

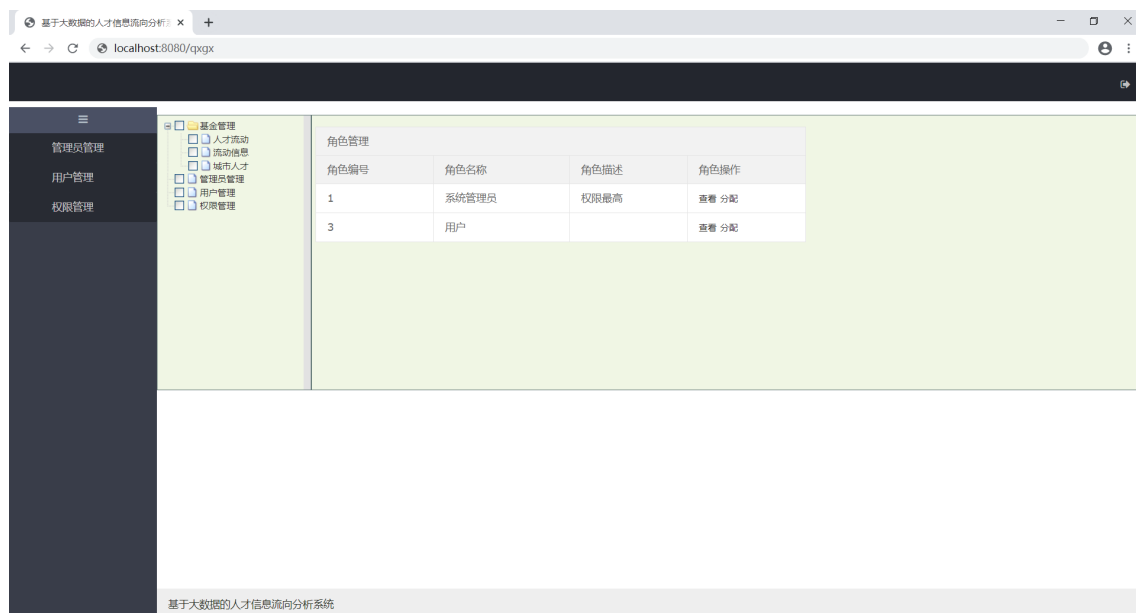


Figure 8. Permission management interface diagram  
图 8. 权限管理界面图

## 7. 结论

本文从人才流动科学的基础出发,依据人才的流动主要因素对于人才流动行为进行分类,主要利用了数据挖掘中的 k-近邻算法。最终实现对城市中新进人才的流向预测,同时也使用数据可视化的方式将人才流动的主要因素进行展示,建立了一个基于大数据的人才信息流向分析系统。将人才流动数据通过大数据方式归类、分析,最后进行数据可视化,建立一个人才信息流向分析系统。该系统可实现对人才的流动信息进行精准分析的同时结合地域分布、成长规律、就业趋势等方面,为人才培育、人才吸引、人才政策等提供较好的参考依据。

## 基金项目

徐州工程学院大学生创新创业训练计划项目(xcx2019001)。

## 参考文献

- [1] 杨雪,樊沼均. 新生代高学历流动人口的流向选择及影响机制[J]. 人口学刊, 2019(6): 64-77.
- [2] 李波,史江萍,杨弘平,等. UML 2 与 Rose 建模从入门到精通[M]. 北京: 清华大学出版社, 2016: 6-11.
- [3] 杨剑锋,乔佩蕊,李永梅,王宁. 机器学习分类问题及算法研究综述[J]. 统计与决策, 2019(6): 36-40.
- [4] 贺雪梅. web 应用开发中的 SSM 框架设计[J]. 电子世界, 2019(1): 206.
- [5] 谢霖铨,赵楠,徐浩,毕永朋. 基于属性相关性的 KNN 近邻填补算法改进[J]. 江西理工大学学报, 2019, 40(1): 95-101.