Design and Implementation of College Students Employment Information Service Platform Based on Data Grasping Technology

Zhengyi Sun[1,a], Jie Dong[1, b] *

[1] No.9, Hunnan East Road, Hunnan District, Shenyang City, Liaoning, P.R.China

[a] 2854835069@qq.com, [b]623236930@qq.com

Abstract. With the expansion of colleges and universities, the number of college graduates has increased year after year, and the employment situation faced by college students is becoming more and more severe. This paper designs and develops an efficient, concise and cost-effective employment information service platform for college students. It uses web crawlers to capture information from multiple angles of recruitment, provides various information such as recruitment information and application for college students, and builds employment information for colleges and universities. The effective sharing of resources, carry forward the advantages of the Internet and e-commerce, achieve the perfect combination of "internet + employment", and provide a broader platform for the employment of college students.

Introduction

With the reform of the employment system, practitioners are increasingly aware that information is the foundation of career choice and the bridge to employers. Whoever gets the information gets the initiative. As the most important feature of this era, the internet has redefined many areas of human life.

There are many similar job search websites on the Internet, such as Zhilian, and Qiancheng. However, the target population of such employment job search websites is oriented to the whole society[1]. The amount of information is huge and cumbersome, and it is difficult for job seekers to efficiently obtain the effective information they really need in a short period of time.

There are two main types of recruitment information for contemporary college students, campus recruitment and social recruitment. However, due to the imbalance of school recruitment information obtained by major universities, there is a phenomenon of information asymmetry. There are also many ways to publish employment information within a school, but there is no platform to integrate employment information. In response to this problem, this paper designs and implements an employment information recommendation system based on web crawler technology. Users can register and log in to the system to search for positions of interest to them, and the system also analyzes the user's information, intelligently matches and recommends the appropriate positions to the users, and also displays the contents of the members' daily browsing. Mark and personalize recommendations to users.

This paper first introduces the basic concepts of web crawler and data cleaning. Then it introduces the overall architecture design of the system, and then elaborates the specific design and implementation of the system.

Web Crawler and Data Cleaning

Web crawler. In order to solve the problem of how to efficiently and conveniently extract and use effective information in massive information, web crawling technology is an important technology that cannot be ignored.

Web crawlers, often referred to as crawlers, are an important part of search engines. A web crawler is a program or script that automatically crawls web information according to certain rules. They are widely used in Internet search engines or other similar websites, and can automatically

collect all the page content they can access to obtain or update the content and retrieval methods of these websites. All web pages grasped by the crawler will be stored by the system, analyzed, filtered, and indexed for later query and retrieval. The Python language is easy to use, and the off-the-shelf crawler framework and toolkit reduce the barriers to use. The use of regular expressions in specific use makes the data capture work vivid and interesting.

Data Cleaning. Data cleaning is the process of re-examining and verifying data to remove duplicate information, correct existing errors, and provide data consistency. Data cleaning is an indispensable part of the process of web crawling, data analysis, etc. And the results directly affect the overall model effect and final conclusions.

Data cleaning is mainly carried out in the following stages: data preprocessing, missing value cleaning, format content cleaning, logical error cleaning, non-demand data cleaning and correlation verification.

System Architecture Design

College student employment information service platform base on the actual situation of the employment work of the school, Combining the actual needs of the counselor and the employment manager in the employment work, through the demand research, sorting out the business process, and referring to the employment platform of other institutions, completing the demand analysis, so as to design and develop the employment information service platform for college students. The platform is closely integrated with the employment management services of college students, guided by information standardization, unified database, unified authentication based，data flow and business flow oriented, so that the system is more in line with the needs of customer business.

College student employment information service platform adopts the Web structure mode. Users inside and outside the school access the system through a web browser, which improves the applicability of the system. The management system is developed and deployed based on the Java EE four-tier technology architecture, providing good cross-platform performance. The client layer components mainly include a web client, an applet, etc., which are located on the client machine. The web layer and the business layer are located on the application server, and the client layer, the network layer, and the business layer access the database in the EIS layer through a unified database connection mechanism. The overall architecture of the system is shown in Figure 1.
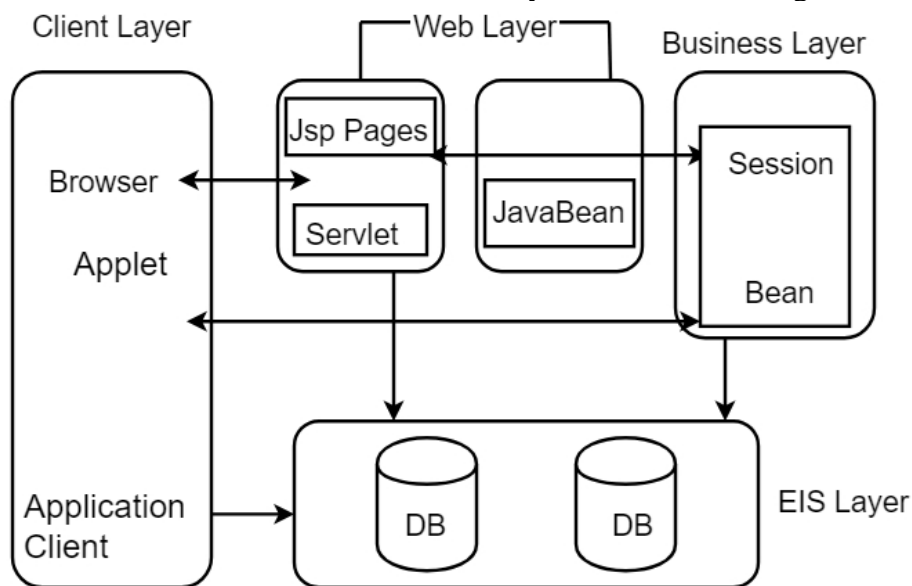

Figure 1.　System architecture diagram

Client Layer. The client layer is the interface logic between the user and the system and is used to run a web browser on the user's machine. Its function is to display data, users can enter data or

update data.

Web Layer. The web layer is responsible for front-end display and user request processing. Its main function is to process user input. It sends the input information to the Enterprise Beans in the business layer for processing.

Business Layer. The business layer concentrates on the main business logic processing in the system and is the core part of the application system. It is mainly responsible for processing the data provided by the client program and sending it to the EIS layer for storage in a timely manner. This process can also be reversed.

EIS Layer. Handling enterprise information system software including database systems[2].

System Design and Implementation

The employment information service platform for college students integrates and analyzes the current employment situation, salary level, development prospects, work environment and other aspects through the internet. This service is mainly aimed at college students who have difficulties in choosing a job, and collects employment information for college students. The selected big data will be intelligently screened and matched with the personal information of college students, and strive to recommend the most suitable professional positions for college students, reduce employment pressure, and clarify the future job search direction.

Project Process. The business of the entire college employment information service platform is divided into three parts: recruitment enterprise, counselor and student user[3]. The specific flow chart is shown in Figure 2.
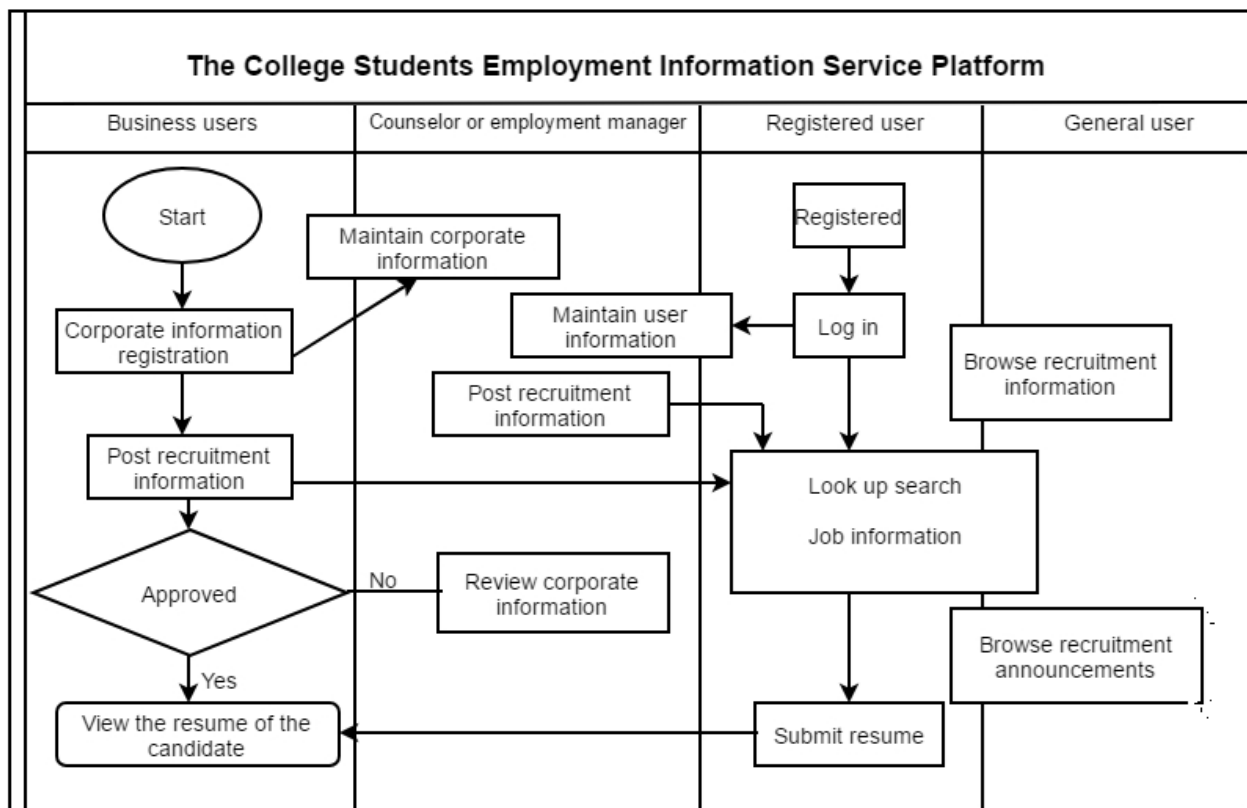

Figure 2   project flow chart

Students can register and log in to the system, and can maintain personal contact information, special hobbies, family status, employment status, etc. through the personal information maintenance module. You can upload multiple resumes and delete your resume. Resumes can be delivered when users browse to their favorite positions[4].

After the enterprise user is successfully registered, the information of the maintenance enterprise

can be maintained, and the job information is uploaded. After the school counselor or the employment manager reviews, the enterprise information and the job information of the enterprise can be displayed on the front page. When a student submits a resume to a position, the business user can view the student's resume or view the student's information.

After registering with the company, the counselor or employment manager needs to review the company's qualifications and corporate security, decide when to recommend the recruitment information to the user, and regularly update the workplace information and the school recruitment announcement to ensure the timeliness of the employment information.

System Function Structure. The main functions of this website include browsing positions, searching for positions, posting resumes, browsing workplace information, paying columns and private ordering. In addition to basic employment information, job recommendation and back office management modules are also available. The job recommendation is that after the user registers and logs in, the system sorts the information registered by the user, and then personally recommends the position to the user. The background management module is designed for platform administrators, enterprise users, and individual users according to different user rights, and adds, deletes, changes, and checks related information[5]. Due to the huge amount of data, a pagination display will be used to facilitate administrators to manage data.

Main Function Design and Implementation. In order to maximize the recruitment information, this website has designed a job capture module to use the web crawler to obtain the design of job information. We use the BeautifulSoup parser to parse the URL text information, analyze the HTML text and page rules of the web page, and then make the following steps to capture the target data: First take the first page code and the last page code，loop through the list page information; after that, the header information is extracted through the th tag, and the information of each row is added to the year and month attributes, and the extracted information is combined into a dataframe format; Finally, the data is cleaned and filtered by Python, and the obtained job information is stored in the database after parsing.

The system maintenance function is the basic operation of adding, deleting, and changing data using AJAX.AJAX is "Asynchronous JavaScript and XML", a web development technology for creating interactive web applications that describes the transformation from web-based web applications to data-based applications. In a data-based application, data required by the user can be obtained from a server independent of the actual web page and can be dynamically written into the web page.

The implementation of the user recommendation function is divided into five parts: When the user does not fill in the job-seeking willingness, the system analyzes other information filled by the user, such as: hobbies, orientations, professional expertise, and other personal subjective information; The naive Bayes classifier is implemented in Python. The user's personal information data is first digitized, and the classifier is used to associate users with the same "feature". In this way, the system can recommend the associated position of the same type of user to the user even if the user does not fill in the job intention. If the user fills in the job-seeking willingness, when there is a corresponding job recruitment, the user can directly receive the recommendation information according to the job-seeking intention. In addition, when the registered user clicks on the browsing system page, the system will determine the intention of the job search according to the user's browsing history and click preferences, and give the recruitment information recommendation. At the same time, the system will also have hot job recommendation and latest job recommendation information.

Provide users with job search functions for different regions and different job categories. In order to avoid the problem of too many places to query, we use JavaScript in the page setup pop-up box for location selection, using Ajax's no-refresh technology to quickly update page resources. To provide users with a good experience, all queries are based on fuzzy query technology, search the database and recommend it to users. For example, if the position filled in by the user is a lawyer and the place of work is Beijing, the system will fuzzy query all the positions in the database that are lawyers and work in Beijing.

Conclusion

The employment information service platform for college students mainly solves the problem of employment difficulties of college students by accurately positioning the employment needs of college students, real-time online recommendation of users and intelligent personalized recommendation, and filling the gap in the employment information consolidation platform in the job market. It is believed that the system will gradually develop into a lubricant for the social system, effectively alleviating the employment pressure of college students and becoming a guiding light for college students' employment.

Acknowledgements

References

[1]G.B. Huang,L.Sun,J.L.Huang,F.M.Yu,L.P.Jia: Neijiang Technology,Vol.29(2018)No.1, p.59. (In Chinese)

[2]T.W.Ni,J.Z.Lin,R.Z.Peng: Journal of Chifeng University,Vol. 33(2017) No.6,p.7. (In Chinese)

[3]D.J.Zheng *Design and Implementation of Information Management System for Graduates in Higher Vocational Colleges*(MS, Jiangsu University, China 2012), p.50. (In Chinese)

[4]C.X. Tan:Design and Implementation of Employment Information Management System in Guangxi Teachers College(MS, University of Electronic Science and Technology, China 2014), p.43. (In Chinese)

[5]F.Hou:Journal of Xi'an Aeronautical University ,Vol. 34(2016) No.1,p.75. (In Chinese)