

Research Article

An Efficient Interpretable Visualization Method of Multidimensional Structural Data Matching Based on Job Seekers and Positions

Guoliang Si ¹, Hengyi Lv ¹, Hangfei Yuan,¹ Dan Xie ² and Ce Peng²

¹Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Jilin 130000, China

²Department of Assets Logistic Management, Jilin University, Changchun 130000, China

Correspondence should be addressed to Dan Xie; xiedan@jlu.edu.cn

Received 26 October 2021; Revised 16 November 2021; Accepted 18 November 2021; Published 24 December 2021

Academic Editor: Gengxin Sun

Copyright © 2021 Guoliang Si et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of Internet technology, millions of small, medium, and microenterprises are using Internet recruitment platforms to host their recruitment information. They have different job requirements and benefits positions. It is important to understand them for job seekers when choosing a position. Existing Internet recruitment platforms do not provide a detailed analysis of positions and visual methods for multidimensional matching of positions and job applicants. Candidates need to spend a lot of energy to screen out suitable positions. In this paper, we propose an efficient interpretable visualization method of multidimensional structural data matching based on job seekers and positions. First, we extract the keywords of the job seeker's ability and benefits based on personal information, and we generate a job seeker ability table and a job seeker demand table. After that, we calculate the degree of the support, confidence, and promotion of each rule through the association rules generated by each frequent itemset of recruitment data to obtain the association rule table. We further explore the relationship between the skills required for the three types of positions based on the association rule. Finally, we use the regression method to build a salary forecasting model. On this basis, we predict the salary of job seekers based on the work experience, education, and work city provided by the job seeker. Simulation results show that our method has better performance on the job analysis and recommendation.

1. Introduction

With the continuous development of society, the competition among enterprises has intensified, and the competition for talents among enterprises has become increasingly fierce. How to find the talents they need more effectively and quickly has become a difficult problem faced by human resource managers of various enterprises. Recruitment websites have gradually become the main way of applying for talents around the world due to their advantages such as large information capacity, low recruitment costs, and breaking time and space constraints [1]. With the development of big data, the data generated by recruitment websites have shown explosive growth. Massive data contain an inestimable value, which mainly includes the demand for market talents, the cultivation of

college students, and the future development direction of various disciplines [2].

To solve the above problems, many domestic and foreign scholars have actively explored and tried. By comparing the research work of various types of recruitment data analysis, from the perspective of recruitment data analysis, existing work can be divided into the following three categories:

- (1) Analyze based on the information of job applicants on the recruitment website. Daniel et al. [3] focused on the study of the relationship between the fit of job seekers using the recruitment website and the satisfaction of the website and their opinions on the organization. They established a theoretical model by using 199 questionnaire surveys and structural equation models based on the recruitment website.

The final results show that the content of the website, the job applicant's understanding of the organization, and the job applicant's social network's recognition of the organization and work in a meaningful way knowing has a positive impact. Zhang and Zheng [4] used text mining techniques to observe the degree of gender differences in expected wages and job-related self-perceptions from recruitment websites, which triggered that the differences in global women's self-perceptions may be related to the belief change theory. They also put forward self-views to explain the expected salary. Wadhawan and Sinha [5] study and identify the factors affecting online job search among students of Delhi from state and private universities using four statistical tools. The final result identified various factors affecting the perception of millennials and postmillennial towards online recruitment and the significant difference among job seekers of different age groups towards different factors. Priyadarshini et al. [6] discussed the influence dimensions, mechanisms, and conditions of the information characteristics of corporate recruitment websites on job seekers' searching and recommendation intentions. Using the stimulus-organism-response framework, the data of 181 job applicants were analyzed. The final result showed that the influence of information characteristics such as information relevance, information accuracy, and information timeliness was determined by the attitude-mediated direct creation of job search intentions and recommendation intentions. These works focus on the analysis and integration of job seekers' information on recruitment websites, ignoring the important impact of enterprise information on data matching.

- (2) Analyze based on corporate information on recruitment websites. Eger et al. [7] found that, according to a survey on recruitment websites, respondents who followed the recruitment information on the company's social media profile did not consider the company's personal information on social networking sites, they believe that the organization's personal information on social networking sites is very important. Baum and Kabst [8] compared the influence of recruitment companies and recruitment websites on the attractiveness of job seekers through research and explained how recruitment activities interact. The final result shows that the influence of companies on the attractiveness of job applicants on the website is significantly stronger than that of print ads. Walker et al. [9] studied the possibility that the diverse clues of companies on the recruitment website affect the processing of the presented information by the website visitors. Using a controlled experiment and a hypothetical organization, it was finally found that, after taking into account the perceived organizational attributes and website design effects, black and

white participants had similar relationships when browsing real organizations and recruiting companies. These works are analyzed based on corporate information on recruitment websites. Although the company's job advertisements have an indirect impact on the attractiveness of job seekers which is mediated by the employer's knowledge, they ignore the impact of multiple recruitment activities on the accuracy of job seekers' personal recommendation results.

- (3) Visual analysis based on recruitment website data. Zhang and Wei [10] and others obtained data-related job recruitment information from different recruitment websites, constructed data-type job characteristics lexicon, studied the characteristics of data-type job talent requirements, and visualized the characteristics and obtained the educational and professional requirements of different positions and differences in professional knowledge and similarities in ability and work experience. Kuhn and Shen [11] conducted research and visual analysis of recruitment data for artificial intelligence at home and abroad. The final study showed that the use of big data measurement methods to speculate on the skill requirements of college graduates by enterprises can reveal the era of artificial intelligence. The content and structure of skills that college graduates need to possess provide enlightenment for colleges and universities to deepen teaching reform. Sodhi and Son [12] crawled professional recruitment data of operations research from recruitment websites, then created a skill dictionary and a keyword dictionary, and used the methods of frequency statistics and content analysis to study the various skills of operations research in different positions. There are differences in demand. The above methods only analyzed the information of job seekers or enterprises from the perspective of visualization and did not give efficient strategies to achieve accurate matching between enterprises positions and job seekers.

Different from the above methods, this article proposes a visual method of multidimensional matching of job applicants and positions, through detailed analysis of job skills requirements, occupational salary, and geographic location, and displays them in a visual manner and through job applicants. The information provided is matched with the characteristics of the recruitment position from multiple angles, and the matching results and job applicant ability analysis report are given from both the ability and demand. Finally, the above method is verified through examples. The main contributions of this paper are shown as follows:

- (1) We extract the keywords of the job seeker's ability and benefits based on personal information, and we generate a job seeker ability table and a job seeker demand table. Base on that, we calculate the degree of the support, confidence, and promotion of each

rule through the association rules generated by each frequent itemset of recruitment data to obtain the association rule table.

- (2) We realize the detailed analysis of job skills requirements, occupational salary, and geographic location, and we display them in a visual manner. We use the information provided by job applicants to match the characteristics of the job position from multiple perspectives and determine the matching results and job applicant ability analysis reports from both capabilities and needs.
- (3) We use the regression method to build a salary forecasting model. On this basis, we predict the salary of job seekers based on the work experience, education, and work city provided by the job seeker. By simulation experiments, we show that the effectiveness of our method on the job analysis and recommendation.

2. Data Acquisition and Analysis

2.1. Data Sources. In this paper, we use the data sets of a job seeker that comes from several major recruitment websites. According to the relevant data, the national online recruitment market reached 10.7 billion yuan in 2019, which has a growth rate of 17.3%. According to the research report on the market development of China's online recruitment industry in 2020, the current recruitment websites are divided into four categories, namely, comprehensive recruitment model, vertical recruitment model, classified information recruitment model, and emerging recruitment model.

The vertical recruitment model is the most efficient recruitment model. It focuses on recruitment services in a certain industry, a specific group of people, or a specific area, which makes it more targeted and professional [13]. Therefore, we select the vertical recruitment model websites as the data sources, which include lagou, liepin, Southern Talent Network, and so on [14]. These data collections include the positions of data analysis, data mining, and software development.

2.2. Data Preprocessing. Among the data collected above, there are some unstructured impurity data on work location and salary. In addition, the recruitment data text contains keywords such as job requirements and job benefits [15]. Therefore, we need to preprocess the recruitment data.

First, we use the method based on the word graph model to obtain the skill keyword set in the introduction of each recruitment information [16]. Then, we segment the job location in the collected recruitment data. Since most of the data we collect are in different cities, we segment the location by keeping the city information. Third, we unified the data format. We use the upper limit of salary and the lower limit of salary to save separately.

2.3. Data Visualization Analysis. In this section, we conduct a detailed visualization analysis of the location of the position, job skill requirements, and the influencing factors of job salary.

- (1) We use the preprocessed three types of data to count the work location of each of the recruitment information. In order to improve the display effect, we set the basic unit of the histogram to compress the number of posts.
- (2) Based on the preprocessed text keywords, we classify the requirements for talents of the companies into six categories: programming language, major, database, personal ability, office software, and professional skills. We make statistics on the above six types of skills and extract high-frequency skill vocabulary from each category. Then, we will save the keyword frequency of the company's capability requirements uniformly that we have counted in excel. The saving format is shown in Table 1.

We consider that there are many types of skills, and there may be some specific associations between the company's skills requirements for talents in each recruitment data. Therefore, we use a dynamic relationship diagram to display them, where different colors are used to indicate the category of skills. We use 1 to indicate the frequency of skills appearing at the same time in each recruitment data using wires. In order to make the relationship between skills more generalized, we recommend setting thresholds to control the connections between skills, which is shown in equation (1), a and b represent two skills required for a position in recruitment data.

$$J_{ab} = \frac{p(a \cap b)}{p(a)} + \frac{p(a \cap b)}{p(b)} = \begin{cases} 1, & J_{ab} > 0.5 \\ 0, & J_{ab} \leq 0.5 \end{cases} \quad (1)$$

- (3) We intercept three types of labels from the data which include education, experience, and salary. First, we process the data hierarchically and store the processed data in excel. The data storage format is shown in Table 2. Second, we use a dynamic mulberry chart to show the relationship between these three labels. The extended branch width in the mulberry chart corresponds to the amount of data flow, which is more suitable for visual analysis of data such as flow. Third, we use each endpoint to segment education, experience, and salary can effectively show the impact of different education and work experience on salary.

3. An Efficient Interpretable Visualization Method of Multidimensional Structural Data Matching

In this section, we propose an efficient Interpretable Visualization method of Multidimensional Structural Data Matching (IVMS-DM), and the specific implementation

TABLE 1: Enterprise capability requirements.

Language	ID						
	345201	354033	156864	614211	941143	162443	235584
Python	1	1	1	1	0	0	1
Java	0	1	1	0	0	0	0
VBA	0	0	0	0	0	0	0
C++	0	0	0	1	1	0	0
C	1	0	0	0	1	0	1
Shell	0	0	0	0	0	1	0
Scala	0	0	0	0	0	0	0
Go	0	0	0	0	0	1	0
MATLAB	0	1	0	0	0	0	0
Javascript	0	1	1	0	1	0	1
Hive	1	0	0	1	0	0	0
Hbase	1	0	0	0	0	0	0
PS	0	0	0	0	0	0	1
Business	0	0	1	1	0	0	0
Logic	1	1	0	1	0	0	0
Math	1	1	0	1	0	0	0
Computer	1	1	0	0	1	0	0
Financial	0	0	1	0	0	0	0

TABLE 2: Information interception of academic experience and salary data.

ID	Source	Target	Value
1	Total number	Bachelor	716
2	Total number	Master	432
3	Total number	Ph.D.	128
4	Total number	Unlimited education	69
5	Total number	Specialist	76
6	Bachelor	Experience 1–3 years	182
7	Bachelor	Experience 3–5 years	389
8	Bachelor	Experience 5–10 years	116
9	Bachelor	Unlimited experience	29
10	Experience 1–3 years	Below 100000 yuan	41
11	Experience 1–3 years	100000 yuan–200000 yuan	52
12	Experience 1–3 years	200000 yuan–300000 yuan	38
13	Experience 1–3 years	300000 yuan–400000 yuan	30
14	Experience 1–3 years	400000 yuan–500000 yuan	21

process shows in Figure 1. We first divide the job seekers according to their abilities, and we judge their work directions using the discrimination model. Then, we introduce a matching algorithm based on the job applicant ability table. Finally, we get the visualized graphs on job distribution and skills relation.

3.1. Discrimination Model. In this subsection, we introduce a discriminant model that ensures the job seeker’s required position highly matches the employees required for the recruitment position. We first divide the personal information provided by the job seekers according to the job seeker’s ability keywords and welfare position keywords are obtained. Then, we generate the job seeker ability table and the job seeker demand table at the same time. We find all frequent itemsets in the database that meet the minimum support threshold based on the recruitment data and

generate the association rules. We calculate the support, confidence, and lift of each rule and generate the association rule table, and we further explore the relationship between the skills required for the three types of positions. We filter the actual situation by setting the lower limit of lift and support. Comparing the job seeker’s ability table with the skill association rule table, the result of judging the job direction of the job seeker is finally obtained.

3.2. Abilities Extraction and Analysis Based on Job Applicants.

In this subsection, we extract and analyze the demand positions based on the abilities of the job applicants. We use U_i to represent the i^{th} job seeker. For each job seeker U_i , we use S_i to represent the set that the skills mastered by U_i , where $S_i = [S_i^1, S_i^2, \dots, S_i^j, \dots, S_i^n]$. We use S_i^j to represent the j^{th} skill of job seeker U_i . In the actual recruitment process of the company, different skills of a job seeker have different effects on whether the job seeker can pass smoothly. For example, data analysis companies want job seekers to master python and hive, and data mining companies want job seekers to master Hadoop and spark. Therefore, in order to describe the role of job seekers’ various skills in job searching accurately, we introduce the weight coefficient W_i , where $W_i = [W_i^1, W_i^2, \dots, W_i^j, \dots, W_i^n]$. As shown in equation (2), a_i^j and d_i^j are the applicant information of U_i in the job seeker ability table A and the job seeker demand table D.

$$W_i^j = a_i^j \frac{P(a_i^j \cap d_i^j)}{S \cdot P(a_i^j) \cdot P(d_i^j)}. \quad (2)$$

According to the value of the weight coefficient, we calculate the similarity of the recruitment positions based on the job skill requirements and the skills of the job seeker, and we add the weight of each skill to the matching results. We sort the final results according to the matching degree by descending.

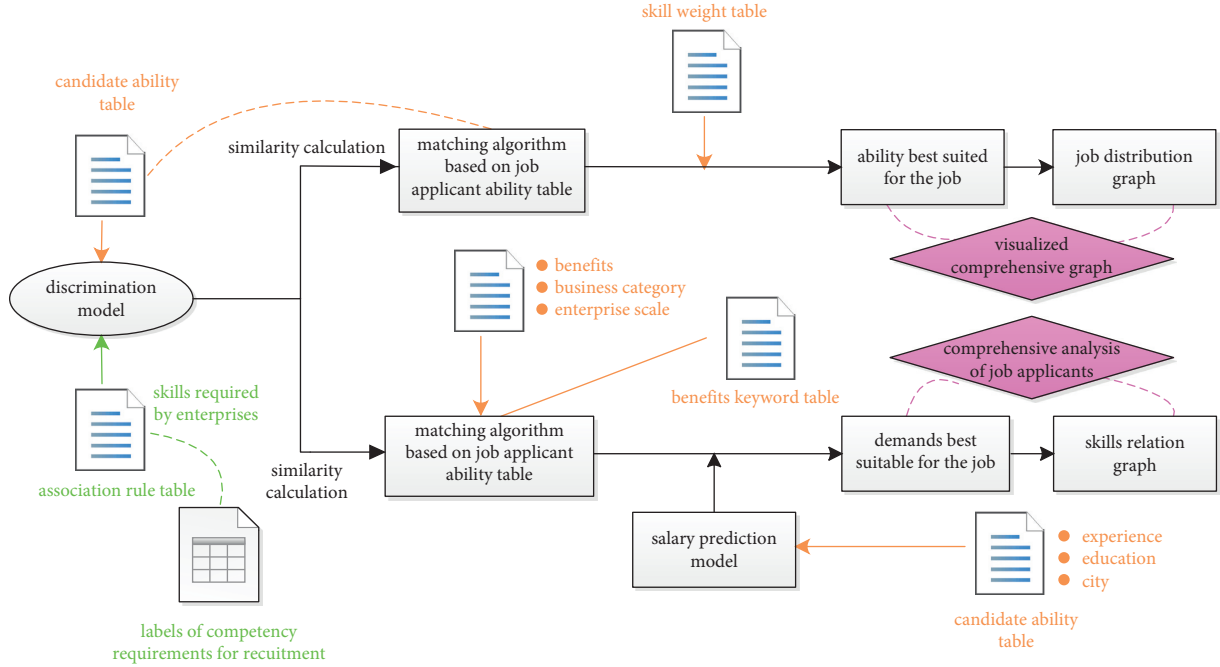


FIGURE 1: The specific implementation flow of IVMS-DM.

3.3. Demands Extraction and Analysis Based on Job Applicants. In this subsection, we propose a salary prediction model based on the regression method. We forecast the salary of job seekers according to the work experience, education background, and work city provided by the job seekers. We extract the welfare-related feature values in the recruitment information text, and we convert the company welfare-related words into a word frequency matrix. Considering that some words appear frequently in the text but are not important, that is, words with a small amount of information need to reduce their weight. Therefore, we use the term frequency-inverse document frequency (TF-IDF) algorithm to convert words related to corporate welfare into vectors that can be processed, and we use IDF values to grasp the importance of words. For a certain keyword t , its TF value calculation formula is shown in (3), where m_{ij} is the frequency of the keyword appearing in the text set. The denominator is the sum of the appearance times of all words in the text set after the word segmentation. The IDF value of the antidocument frequency is shown in (4), where $|\{i: t_j \in h_j\}|$ represents the number of texts containing the keyword. We finally calculate the TF-IDF value of the keyword and generate a graph of corporate welfare benefits.

$$TF_{i,j} = \frac{m_{ij}}{\sum_{j=1}^k m_{ij}}, \quad (3)$$

$$IDF_j = \log \frac{|H|}{|\{i: t_j \in h_j\}|}. \quad (4)$$

Furthermore, we calculate the similarity using the job seeker demand table and the relevant feature values of corporate welfare in the recruitment data. Then, the results are sorted according to the similarity of the demands of job

seekers' welfare. We filter out the positions whose salary is greater than the salary predicted by the job seeker, which is the final matching result.

3.4. Display Job Matching Analysis Report. According to the job ID matched by the job seeker's demands, the skills needed for the jobs are obtained from the database. Then we calculate the proportion of each skill in each city. We use a radar chart to display the required skills for job seekers. According to the ideal city of the job seeker, we divide the radar chart and use equation (5) to calculate the probability of finding a job that meets the seeker's demands in each ideal city. The suggestions based on the above information are given to job seekers. SUM is the total number of positions, and is shown in equation (6). Pos and Neg represent two matching results: Pos is the position based on the needs of job seekers according to the intention of job seekers, and Neg is the position based on the abilities of job seekers according to the needs of enterprises.

$$\text{probability} = \left[\frac{\sum_{i=0}^n (P_i/\text{Pos}) * W_i}{i} + \frac{\sum_{j=0}^n (P_j/\text{Neg}) * W_j}{j} \right], \quad (5)$$

$$\text{SUM} = \text{Pos} + \text{Neg}. \quad (6)$$

4. Simulation and Result

4.1. Simulation Setting. We download the dataset from a recruitment website [17]. This dataset contains 14936 recruitment data in 20 days that mainly including data analysis, data mining and software development. We select 13273 valid data by preprocessing, and the experimental part uses Python and related libraries.

4.2. Recruitment Data Visualization

4.2.1. Analysis of Position Location. We analyze the distribution of enterprises across the country, and we set the basic unit as 50. The national distribution map of enterprises is obtained in Figure 2. From the perspective of cities where enterprises are located, Beijing, Shenzhen, and Shanghai are the top three cities in the 660 cities in China which are accounting for 34%, 26%, and 18% respectively. From the perspective of regional distribution of enterprises, enterprises in the Yangtze River Delta and Guangdong Province are more widely distributed which are accounting for 36% and 28% respectively. From the perspective of enterprise-scale, listed companies with more than 2000 people accounted for the largest proportion. These companies are mainly distributed in Beijing and Shanghai, which are accounting for 17% and 15% respectively. From the perspective of the enterprises' industry, data services and mobile Internet accounted for the largest proportion, which are accounting for 38% and 32%, respectively.

4.2.2. Analysis of Post Skill Correlation. The association of the post skills is shown in Figure 3. According to the results, we have the following observations: (i) The top three discipline requirements are computer, mathematics, and statistics, and these three positions have similar discipline requirements for employees. In terms of the requirements of the abilities, the three positions also value personal learning ability and communication ability. (ii) There are differences in the importance of different skills for different positions. We can see that the direction of computer science needs to have a solid mathematical foundation, and it also needs to have a strong ability to receive new things. Most enterprises in the data analysis direction need to master python, hive, and SQL at the same time, while enterprises in the data mining direction need to master Python, Hadoop, and spark at the same time. In terms of the programming language of the three positions, Java ranks in the top three. However, MySQL accounts for a large proportion in terms of database.

4.2.3. Analysis of the Multidimensional Influencing Factors of Salary. In this subsection, we analyze the multi-dimensional influencing factors of salary of three types of positions which are shown in Figures 4–6. From the perspective of educational requirements of enterprises, the proportion of undergraduates in the three types of positions is the largest that accounting for 58%, 43%, and 72%, respectively. Among them, the number of enterprises requiring a master's degree in data mining is the largest. What's more, the software development direction requires the number of junior college students to rank second. From the above data, we can see that the software development direction has the lowest education requirements, while the data mining direction has the highest education requirements; From the perspective of experience requirements of enterprises, these three types of positions are generally concentrated in 3–5 years, and the undergraduate education accounts for the largest proportion. Most enterprises generally require 1–3 years of personnel experience for a master degree, but there is generally no experience limitation. From the perspective of salary given by enterprises, the average salary of data mining is the highest, with an annual salary of 0.3-0.4 million, followed by that of data analysis of 0.2-0.3 million, and that of software development of 0.1-0.2 million. From the relationship among education background, experience, and salary, the proportion of undergraduates working in data analysis with an annual salary of 0.2-0.3 million for 3–5 years is the highest. The number of software development graduates with 1–3 years' salary of 0.1-0.2 million is the largest, and the number of data mining graduates with 1–3 years' salary of 0.4-0.5 million is the largest.

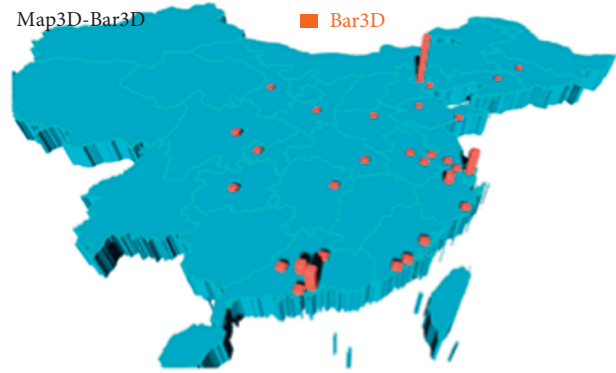


FIGURE 2: National distribution map of enterprises.

proportion. Most enterprises generally require 1–3 years of personnel experience for a master degree, but there is generally no experience limitation. From the perspective of salary given by enterprises, the average salary of data mining is the highest, with an annual salary of 0.3-0.4 million, followed by that of data analysis of 0.2-0.3 million, and that of software development of 0.1-0.2 million. From the relationship among education background, experience, and salary, the proportion of undergraduates working in data analysis with an annual salary of 0.2-0.3 million for 3–5 years is the highest. The number of software development graduates with 1–3 years' salary of 0.1-0.2 million is the largest, and the number of data mining graduates with 1–3 years' salary of 0.4-0.5 million is the largest. From the perspective of educational requirements of enterprises, the proportion of undergraduates in the three types of positions is the largest that accounting for 58%, 43%, and 72%, respectively. Among them, the number of enterprises requiring a master degree in data mining is the largest. What's more, the software development direction requires the number of junior college students to rank second. From the above data, we can see that the software development direction has the lowest education requirements, while the data mining direction has the highest education requirements; From the perspective of experience requirements of enterprises, these three types of positions are generally concentrated in 3–5 years, and the undergraduate education accounts for the largest proportion. Most enterprises generally require 1–3 years of personnel experience for a master degree, but there is generally no experience limitation. From the perspective of salary given by enterprises, the average salary of data mining is the highest, with an annual salary of 0.3-0.4 million, followed by that of data analysis of 0.2-0.3 million, and that of software development of 0.1-0.2 million. From the relationship among education background, experience, and salary, the proportion of undergraduates working in data analysis with an annual salary of 0.2-0.3 million for 3–5 years is the highest. The number of software development graduates with 1–3 years' salary of 0.1-0.2 million is the largest, and the number of data mining graduates with 1–3 years' salary of 0.4-0.5 million is the largest.



FIGURE 3: The association of the post skills.

4.3. Post Matching and Visualization

4.3.1. *Division of the Job Seekers' Information.* The job seeker's personal information is divided into the job seeker's ability table and the job seeker's demand table, and the split result is shown in Tables 3 and 4.

4.3.2. *Discrimination of Job Seeker's Work Direction.* We set the lower limit of life as 1, and the lower limit of support as 0.2. We use the processed data to make the skill and position association rules table and calculate the similarity between the candidate's ability table and the position skill association

rules table. The result shows that the candidate's direction is data analysis. The table of skill post association rules is shown in Table 5.

4.3.3. *Job Matching Based on Job Seeker Competency Table.* We calculate the weight of each skill and construct the coefficient table, which is shown in Table 6. Then, we calculate similarities of the skills and divide the recruitment positions according to the required skills. Then, the ability table of job seekers is used to match the divided positions. In the process of matching, the coefficient is added to the algorithm, and finally the most suitable position set is generated according to

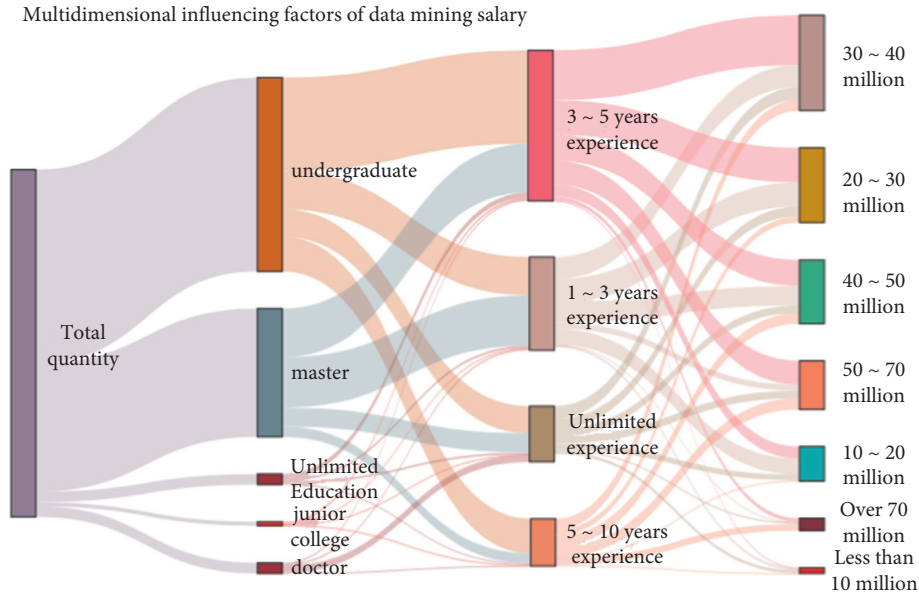


FIGURE 4: Multidimensional influencing factors of data mining salary.

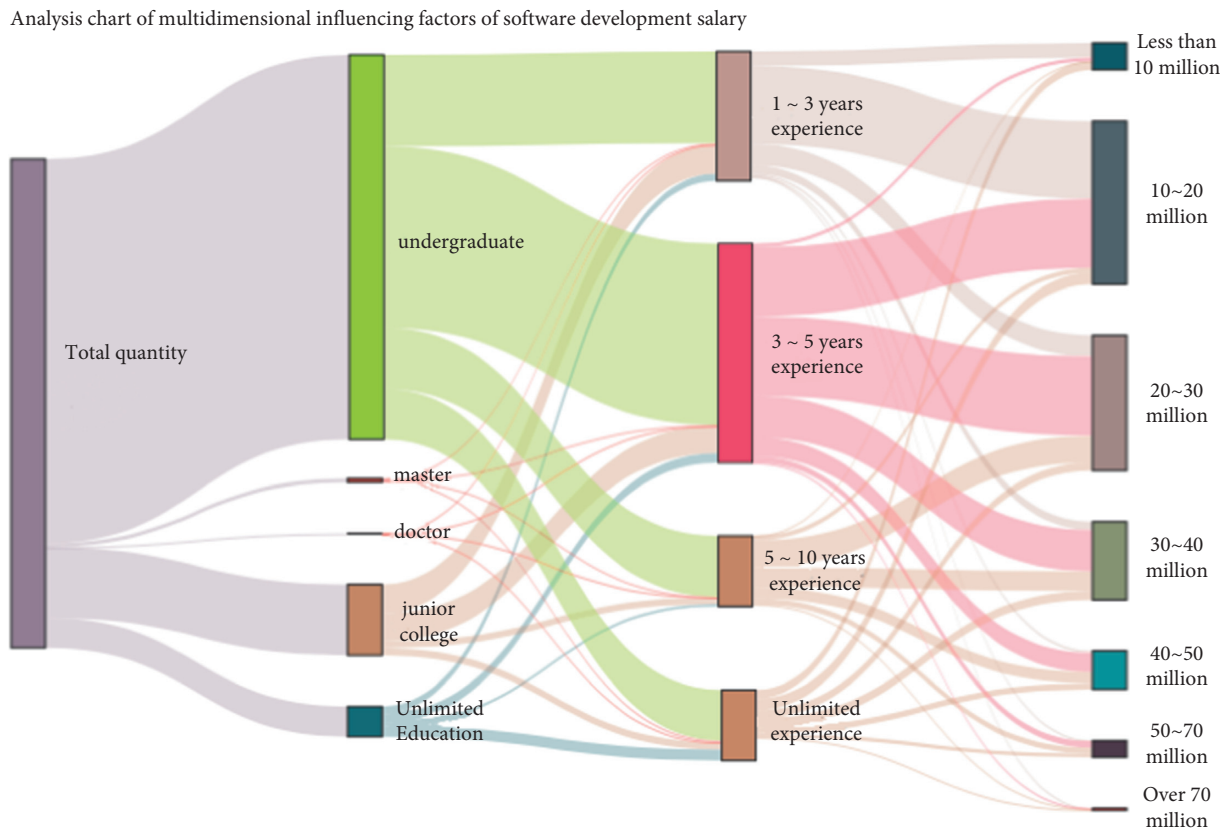


FIGURE 5: Analysis chart of multidimensional influencing factors of software development salary.

the descending order of the matching degree. The result contains 258 pieces of recruitment information, and the most suitable position set is shown in Table 7.

4.3.4. Job Matching Based on Job Seeker Demand Table. According to the regression model, the annual salary of job seekers is 213750 yuan.

Data analysis multidimensional influencing factor analysis chart of salary

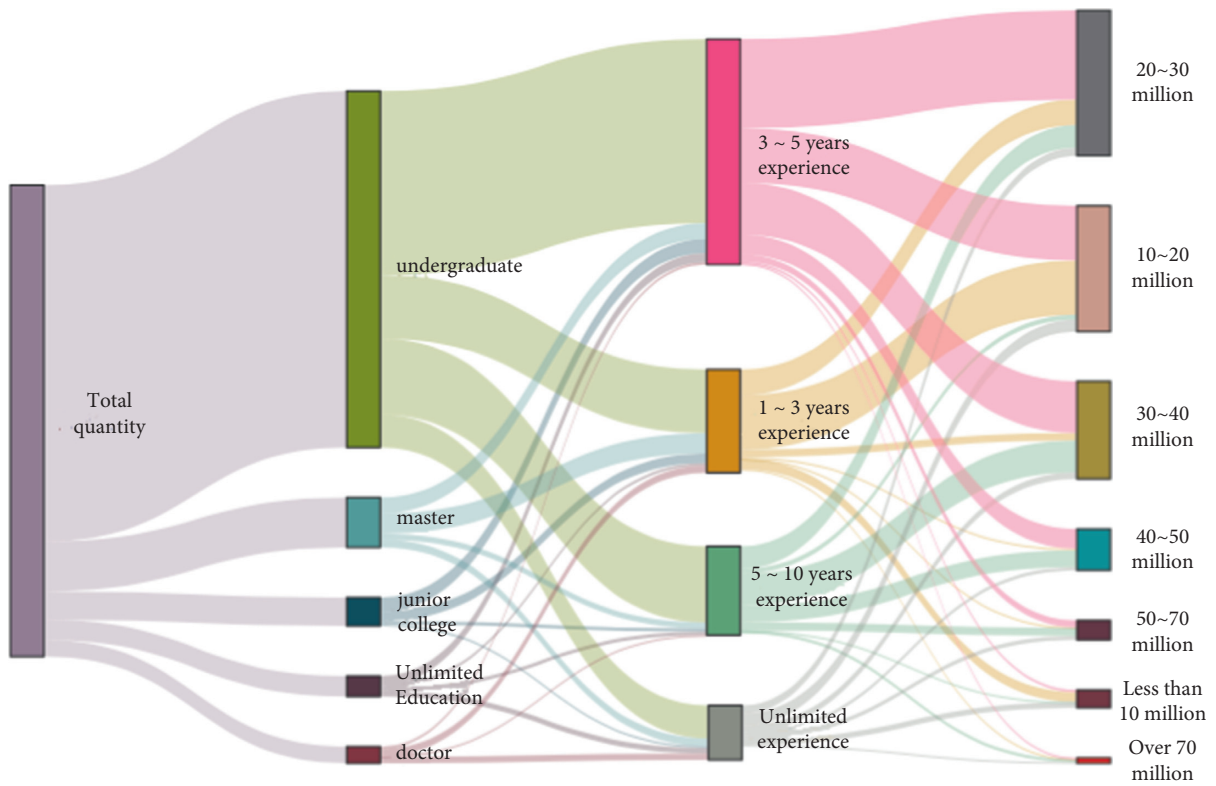


FIGURE 6: Data analysis multidimensional influencing factor analysis chart of salary.

TABLE 3: Job seeker’s ability table.

Programing language	Python, Java
Database	MySQL
Office skills	Excel
Professional skills	Spark, html, CSS
Personal ability	Learning ability, logical ability, and professional ability
Learning direction	Computer
Education	Undergraduate
Working experience	3 years

TABLE 4: Job seeker’s demand table.

Ideal city	Beijing, Shanghai, Hangzhou
Welfare	Insurances, meals free
Scale	Large
Category	Internet, e-commerce, entertainment

TABLE 5: Skill and position category association rules table.

	<i>Antecedents</i>	<i>Consequents</i>	<i>Antecedent support</i>	<i>Consequent support</i>
38	Mathematics	Statistics	0.408134	0.12923
52	Hadoop	Hive	0.219617	0.141148
54	Spark	Hive	0.255502	0.141148
56	Hadoop	Spark	0.219617	0.255502
84	Spark, Python	Java	0.192344	0.263158
143	Hadoop, Python	Spark	0.166986	0.255502
218	Computer, Hadoop	Spark	0.142105	0.255502

TABLE 5: Continued.

240	Hadoop, spark	Hive	0.197608	0.141148	
246	Machine learning, Hadoop	Spark	0.169856	0.255502	
365	Machine learning, Hadoop, Python	Spark	0.13445	0.255502	
404	Hadoop	Machine learning, computer, spark	0.219617	0.141627	
405	Spark	Machine learning, computer, Hadoop	0.255502	0.115789	
40	Spark	Hadoop	0.282773	0.268067	
43	HTML	CSS	0.12605	0.138655	
80	Spark, Python	Hadoop	0.157143	0.268067	
82	Hadoop, Python	Spark	0.143277	0.282773	
88	Hive, Java	Spark	0.114286	0.282773	
98	CSS, Java	HTML	0.120168	0.12605	
122	Spark, computer	Hadoop	0.165546	0.268067	
128	Spark, Hive	Hadoop	0.195798	0.268067	
42	C++, computer	HTML	0.219617	0.141627	
48	Computer, Android	Java	0.255502	0.115789	
	<i>Support</i>	<i>Confidence</i>	<i>Lift</i>	<i>Conviction</i>	<i>Family</i>
38	0.107177	0.262603	2.119071	1.188066	Data mining
52	0.118182	0.538126	3.812488	1.859495	Data mining
54	0.124402	0.486891	3.449502	1.67382	Data mining
56	0.197608	0.899782	3.521619	7.428791	Data mining
84	0.103349	0.537313	2.041791	1.59253	Data mining
143	0.147847	0.885387	3.465278	6.495742	Data mining
218	0.129187	0.909091	3.558052	8.189474	Data mining
240	0.109091	0.552058	3.911191	1.917328	Data mining
246	0.156938	0.923944	3.616184	9.788765	Data mining
365	0.122967	0.914591	3.579578	8.716826	Data mining
404	0.106699	0.485839	3.430416	1.669463	Data mining
405	0.106699	0.417603	3.606571	1.518226	Data mining
40	0.218908	0.774146	2.887879	1.67382	Data analysis
43	0.121849	0.966667	6.971717	5.840336	Data analysis
80	0.118908	0.756684	2.822742	3.008163	Data analysis
82	0.118908	0.829912	2.934904	4.216799	Data analysis
88	0.1	0.875	3.094354	5.737815	Data analysis
98	0.105882	0.881119	6.99021	7.351458	Data analysis
122	0.129412	0.781726	2.916156	3.353273	Data analysis
128	0.153782	0.785408	2.929891	3.410807	Data analysis
42	0.106699	0.485839	3.430416	1.669463	Software development
48	0.106699	0.417603	3.606571	1.518226	Software development

TABLE 6: skill and construct the coefficient table.

<i>Skill</i>	<i>Classification</i>	<i>Weight</i>
Hadoop	Data mining	1.252852
Python	Data analysis	1.837236
Python	Data mining	1.300227
Spark	Data mining	1.6473
Hadoop	Data analysis	1.319256
Java	Software development	2.23734
Hive	Data analysis	2.103453
C++	Software development	1.84724
MySQL	Software development	2.4235
Excel	Data analysis	1.36742
SQL	Software development	1.2253
Spark	Data analysis	1.86534
<i>Skill2</i>	<i>Classification3</i>	<i>Weight4</i>
Hadoop, python	Data mining	2.423652
HTML, MySQL	Software development	1.93652

TABLE 6: Continued.

Java, C++	Software development	1.87234
HTML, CSS	Software development	1.34634
Oracle, C++	Software development	1.765345
Python, Spark	Data mining	1.628372
Linux, C	Software development	1.837236
Python, Hive	Data analysis	1.963872
MySQL, python	Data analysis	1.13267
Python, Excel	Data mining	2.037423
BI, Hive	Data analysis	1.46234
Hadoop, Spark	Data mining	1.432423

TABLE 7: The most suitable position set.

position_id	Title	company_URL	Skill
192375	Web front end	http://www.lagou.com/gongsi/30008.html	CSS, Mobile terminal, Front end development, HTML5
1675817	Senior Java R&D Engineer	http://www.lagou.com/gongsi/1686.html	CSS, Spark, HTML, Java
1763729	Big data development engineer	http://www.lagou.com/gongsi/12597.html	Hadoop, Spark, Hive, Java
2873774	Front end engineer (Beijing/Wuhan)	http://www.lagou.com/gongsi/31164.html	Learning ability, Computer direction, HTML, Javascript, Web front end
2976934	Big data development engineer	http://www.lagou.com/gongsi/189298.html	Excel, Spark, MySQL, Java, Data analysis
...
7459738	Big data development engineer	http://www.lagou.com/gongsi/120642419.html	Spark, python, Algorithm Engineer
7459835	Front end expert	http://www.lagou.com/gongsi/153849.html	Java, Logical ability, Computer direction, Excel, MySQL, Javascript
7460129	Senior data analyst	http://www.lagou.com/gongsi/211706.html	Spark, python
7460663	Big data development engineer	http://www.lagou.com/gongsi/577862.html	Hadoop, hbase, Spark, python
7460888	Front end senior development engineer	http://www.lagou.com/gongsi/120644511.html	CSS, HTML5, CSS3, jQuery, Javascript

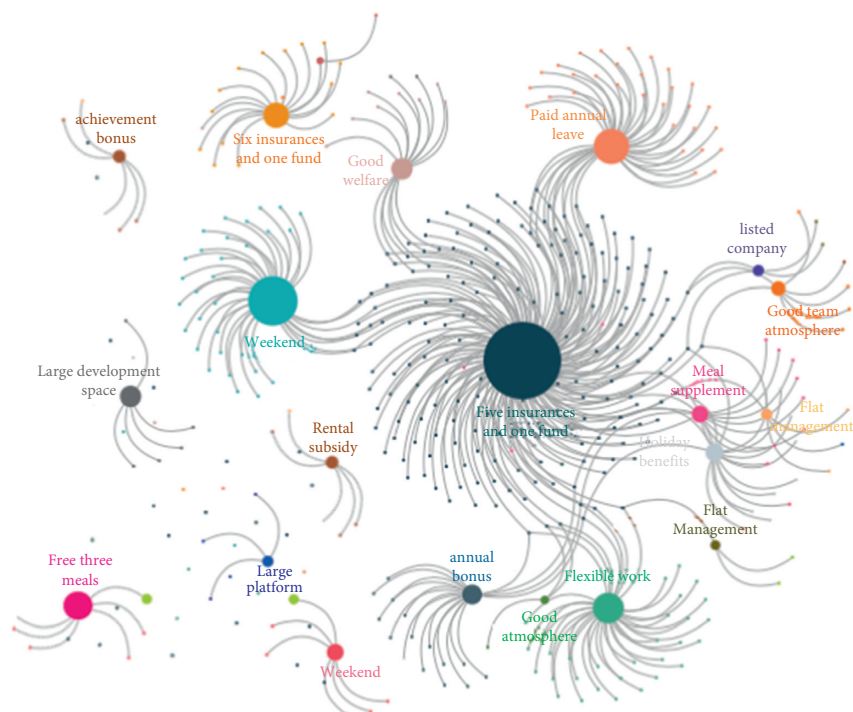


FIGURE 7: The relationship diagram of enterprise welfare.

TABLE 8: The set of positions with the most suitable demand.

<i>position_id</i>	<i>Title</i>	<i>enterprice_locale</i>	<i>enterprice_category</i>
4553174	Commercial data analyst Data visualization R&D	Beijing	Entertainment
5119472	Director - platform direction	Shanghai	Entertainment
5120075	Data Analyst	Beijing	Entertainment
5120103	Data Analyst	Shanghai	Electronic commerce
5120207	Senior Data Analyst	Beijing	Entertainment
...
7377471	Senior Data Analyst	Shanghai	Entertainment
7388404	Big data Development Engineer - live broadcast	Beijing	Mobile Internet
7419188	Big data Development Engineer - Zhu Xiaobang	Beijing	Entertainment
7447755	Senior Data Analyst	Hangzhou	Mobile Internet
7451860	Short video data operation manager	Beijing	Entertainment

<i>position_id</i>	<i>enterprice_stage</i>	<i>enterprice_people</i>	<i>Money</i>	<i>Text</i>
4553174	C	2000	240000 480000	Six insurances and one fund, flexible work, free three meals, rental subsidy, flat management, professional Daniel, promotion opportunity, good team atmosphere
5119472	C	2000	480000 960000	Six insurances and one fund, free three meals, rental subsidy, professional Daniel
5120075	C	2000	180000 360000	Six insurances and one fund, flexible work, free three meals, good team atmosphere
5120103	C	1000	180000 360000	Six insurances and one fund, flexible work, free three meals, good team atmosphere
5120207	C	2000	180000 360000	Six insurances and one fund, flexible work, free three meals, good team atmosphere
...
7377471	C	2000	180000 360000	Six insurances and one fund, stock infringement, flexible work, free three meals
7388404	C	2000	300000 600000	Six insurances and one fund, stock infringement, flexible work, free three meals
7419188	C	2000	240000 480000	Six insurances and one fund, flexible work, free three meals, rental subsidy
7447755	C	1000	240000 360000	Six insurances and one fund, flexible work, free three meals, rental subsidy
7451860	C	2000	180000 360000	Six insurances and one fund, flexible work, free three meals, rental subsidy

After calculation, the relationship diagram of enterprise welfare is shown in Figure 7. The similarity between the job seeker's demand table and the welfare benefits extracted from the recruitment information is calculated, and the predicted salary of the job seeker is filtered. Finally, 63 positions with the most suitable demand are obtained, and the set of positions with the most suitable demand is shown in Table 8.

5. Conclusions

In this paper, we propose a visualization method of multidimensional matching between job seekers and positions. First, we use visualization to analyze the key information of the position in detail and give the analysis results. Then, we match the recruitment positions from two perspectives: one is based on the ability of the job seeker, and the other is based

on the needs of the job seeker. According to the matching results, it provides targeted professional suggestions to job seekers and plays the role of the website as a link between enterprises and job seekers. Finally, we visually analyze the three types of job recruitment data in salary, geographical location, and skills. According to the abilities and demands of job seekers, two comprehensive analysis charts are generated which include post display, multidimensional feature analysis chart and analysis report of job seekers' abilities. The experiment results show that our method can achieve the expected effect. At the same time, there are also shortcomings. The data source of this article has certain limitations. The selected experimental testers do not cover all types of user groups of recruitment websites, which leads to the limited representative of the testers. The research object selection is limited to a kind of comprehensive recruitment website. However, the development of the online

recruitment industry is rapid, and a variety of new online recruitment models have emerged at present. Our future research should cover a variety of domestic and foreign recruitment websites and explore job seekers' satisfaction with the website and the success rate of job seekers in the recruitment website. Only in this way can we study and analyze job seekers and recruitment information in a wider range and on a deeper level.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This paper was supported by the Space camera vibration parameters detection and blurred image restoration by using rolling shutter CMOS autocorrelation imaging (Grant no.: 62005269). Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences.

References

- [1] X. G. Zhou, S. L. Chen, and L. Q. Li, "Research on the relationship among big data capacity, technological innovation and competitiveness of human resource service enterprises," *Management Review*, vol. 33, no. 7, pp. 81–91, 2021.
- [2] A. Aizawa, "An information-theoretic perspective of TF-IDF measures," *Information Processing & Management*, vol. 39, no. 1, pp. 45–65, 2003.
- [3] E. Daniel, R. Stone, and L. Baker-Eveleth, "The influence of recruitment websites on job-seeker perceptions of organization and job fit," *International Journal of Technology and Human Interaction*, vol. 14, no. 4, pp. 1–22, 2018.
- [4] X. Zhang and Y. Zheng, "Gender differences in self-view and desired salaries: a study on online recruitment website users in China," *PLoS One*, vol. 14, no. 1, Article ID e0210072, 2019.
- [5] S. Wadhawan and S. Sinha, "Factors influencing young job seekers perception towards job portals," *AIMS International Journal of Management*, vol. 12, no. 3, pp. 199–212, 2019.
- [6] C. Priyadarshini, S. Sreejesh, and R. R. Jha, "Impact of informational characteristics of the recruitment website on graduating students' job pursuit intention," *International Journal of Human Capital and Information Technology Professionals*, vol. 10, no. 2, pp. 1–19, 2019.
- [7] L. Eger, M. Mičik, and P. Rehoř, "Employer branding on social media and recruitment websites: symbolic traits of an ideal employer," *E+M Ekonomie a Management*, vol. 21, no. 1, pp. 224–237, 2018.
- [8] M. Baum and R. Kabst, "The effectiveness of recruitment advertisements and recruitment websites: indirect and interactive effects on applicant attraction," *Human Resource Management*, vol. 53, no. 3, pp. 353–378, 2014.
- [9] H. J. Walker, H. S. Feild, J. B. Bernerth, and J. B. Becton, "Diversity cues on recruitment websites: investigating the effects on job seekers' information processing," *Journal of Applied Psychology*, vol. 97, no. 1, Article ID 214, 2012.
- [10] J. F. Zhang and R. B. Wei, "Study on the job characteristics of data-type post demand from the recruitment website in China," *Journal of Intelligence*, vol. 37, no. 6, pp. 176–182, 2018.
- [11] P. J. Kuhn and K. Shen, "Do employers prefer undocumented workers? Evidence from China's Hukou system (no. 8289)," IZA Discussion Papers, 2014.
- [12] M. S. Sodhi and B.-G. Son, "Content analysis of OR job advertisements to infer required skills," *Journal of the Operational Research Society*, vol. 61, no. 9, pp. 1315–1327, 2010.
- [13] B. D. Qu, "A research of the employment problem on common job-seekers and graduates," *International Journal of Automation and Computing*, vol. 6, no. 3, Article ID 314, 2009.
- [14] J. P. Wieg, F. Drasgow, and J. Rounds, "Misfit matters: a re-examination of interest fit and job satisfaction," *Journal of Vocational Behavior*, vol. 125, Article ID 103524, 2021.
- [15] K. J. Eschleman, D. Mast, Q. Coppler, and J. Nelson, "Organizational factors related to attracting job seekers higher in hardiness," *International Journal of Selection and Assessment*, vol. 27, no. 2, pp. 169–179, 2019.
- [16] Y. C. Chang, Y. X. Zhang, H. Wang, H. Y. Wan, and C. J. Xiao, "Features oriented survey of state-of-the-art key phrase extraction algorithms," *Ruan Jian Xue Bao/Journal of Software*, vol. 29, no. 7, pp. 2046–2070, 2018.
- [17] <https://www.lagou.com/>.