A Job Recommender System Based on User Clustering

Wenxing Hong, Siting Zheng, Huan Wang*
School of Information Science and Technology, Xiamen University, Xiamen, China
Email: hwx@xmu.edu.cn, zhengsiting@gmail.com, wanghuan@xmu.edu.cn

Jianchao Shi
Department of Computer Science, Westsouth Petrolum University, Chengdu, China
Email: sjcgainhan@126.com

***Abstract*—In this paper, we first provide a comprehensive investigation of four online job recommender systems (JRSs) from four different aspects: user profiling, recommendation strategies, recommendation output, and user feedback. In particular, we summarize the pros and cons of these online JRSs and highlight their differences. We then discuss the challenges in building high-quality JRSs. One main challenge lies on the design of recommendation strategies since different job applicants may have different characteristics. To address the aforementioned challenge, we develop an online JRS, iHR, which groups users into different clusters and employs different recommendation approaches for different user clusters. As a result, iHR has the capability of choosing the appropriate recommendation approaches according to users' characteristics. Empirical results demonstrate the effectiveness of the proposed system.**

***Index Terms*—online job recommender system, user cluster, recommendation approach**

I. Introduction

Recently, job recommendation has attracted a lot of research attention and has played an important role on the online recruiting website. Different from traditional recommendation systems which recommend items to users, job recommender systems (JRSs) recommend one type of users (e.g., job applicants) to another type of users (e.g., recruiters). In particular, job recommender system is designed to retrieve a list of job positions to a job applicant based on his/her preferences or to generate a list of job candidates to a recruiter based on the job requirements. To obtain a good recommendation results, many recommendation approaches are presented and applied in the JRS. Typically, given a user, existing JRSs employ a specific recommendation approach to generate a ranked list of jobs/candidates. However, different users may have different characteristics and a single recommendation approach may not be suitable for all users. Therefore, a high-quality JRS should have the capability of choosing the appropriate recommendation approaches according to the user's characteristic.

In this paper, we develop a local JRS called iHR to address the aforementioned issue in job recommendation. iHR classifies the user into groups by using both the individual information and historical behaviors of users, and then employs the corresponding recommendation approach for each user group.

In summary, the contribution of this paper is threefold:

a) From a product perspective, we highlight the differences of four online JRSs in four areas: user profiling, recommendation strategies, recommendation output, and user feedback. The advantages and disadvantages of these online JRSs are also listed for having a good understanding of existing online JRSs.

b) By comparing with the generic RS, we outline the specific challenges essential to the development of a JRS. The solutions to the challenges are helpful for improving both the accuracy and efficiency of the recommender system (RS).

c) We develop an online JRS capable of choosing the suitable recommendation approach for different user groups for Xiamen talent service center. The user groups are constructed based on their individual information and historical behaviors.

The rest of the paper is organized as follows. Section II presents a literature review about the technical approaches of the JRS. In Section III, four online JRSs are compared and analyzed at the product level. It also describes the differences between a JRS and a generic RS. In Section IV, we develop a novel JRS by clustering the users and finding out the appropriate recommendation approach for each user group. Finally, Section V contains some conclusions plus some ideas for future work.

II. Related Work

The JRS has been studied from many aspects. Al-Otaibi et al. [1] summarized the categories of existing online recruiting platforms and listed the advantages and disadvantages of technical approaches in different JRSs. For example, bidirectional recommendation is accomplished but only binary representation is allowed in the probabilistic hybrid approach. We also had done some research on feature extraction, resume mining, recommendation approach, ranking, and explanation for the JRS. In our previous work [2], user profiling and calculating similarity are presented as the prevailing process of a JRS, and the architecture and product features are briefly discussed. Moreover, empirical experiments had been conducted on a local online

doi:10.4304/jcp.8.8.1960-1967

recruiting website and details on the specific case study are illustrated in Section IV.

From the technical perspective, JRS has been classified into five categories described as follows:

a) Content-based Recommendation (CBR)

The principle of a content-based recommendation is to suggest items that have similar content information to the corresponding users. For example, in the recommendation that recommending jobs to a job applicant, the content is the personal information and their job desires. While recommending candidates to recruiters, the job description posted by recruiters, including the background description of enterprises, are used as the content for recommendation.

The basic process of content-based recommendation is acquiring the content information of job applicants and jobs and calculating their similarities. So the content information plays an important role in the content-based recommendation [3]. Yu et al. [4] presented a cascaded extraction approach for resumes to obtain the more effective information. Yi et al. [5] built a relevance-based language model – Structured Relevance Models for modeling and retrieving semi-structured documents. Furthermore, Paparrizos et al. [6] trained a machine learning model to predict candidates' next job transition based on their past job histories as well as the data of both candidates and enterprises in the web.

b) Collaborative Filtering Recommendation (CFR)

Collaborative filtering recommendation, known as the user-to-user correlation method, finds similar users who have the same taste with the target user and recommends items based on what the similar users like. The key step in CFR is computing the similarities among users. Collaborative filtering recommendation algorithm can be classified into memory-based and model-based [7, 8]. In the memory-based collaborative filtering recommendation, a user-item rating matrix is usually used as the input [9, 10]. Applied in the job recruiting domain, some user behaviors or actions can generate the user-item rating matrix according to the predefined definitions and transition rules. Färber et al. [11] presented an aspect model to produce a rating matrix that assigns assessed values to candidate's profile using the Expectation Maximization (EM) algorithm.

c) Knowledge-based Recommendation (KBR)

In the knowledge-based recommendation, rules and patterns obtained from the functional knowledge of how a specific item meets the requirement of a particular user, are used for recommending items [12]. For example, employees who have one or more years of work experience exhibit better performance as compared to those without experience. This can be used as a job performance rule in the online recruiting. Chien et al. [13] developed a data mining framework based on decision tree and association rules to generate useful rules for selecting personnel feature and enhancing human capital. In addition, other types of knowledge such as ontology can also be used in the job recommendation. Lee and Brusilovsky [14] employed an ontology checker to match information with ontology and perform the classification in the JRS.

d) Reciprocal Recommendation (ReR)

Firstly proposed by Luiz Pizzato et al. [15], reciprocal recommender is a special kind of recommender systems. The preferences of all the users are taken into account and need to be satisfied at the same time. As a result, ReR achieves a win-win situation for users and improves the accuracy of recommender systems that match people and people.

Yu et al. [16] proposed a similarity calculation method for calculating the reciprocal value and achieving the reciprocal recommendation based on the explicit preferences obtained from users' resumes and the implicit preferences acquired from the user's interaction history. Malinowski et al. [17] also used a bilateral recommendation approach which considers the two parts of JRS to match the job applicants and jobs. Li et al. [18] proposed a generalized framework for reciprocal recommendation that is applied to online recruiting, in which they model the correlations among users by a bipartite graph.

e) Hybrid Recommendation (HyR)

All recommendation approaches mentioned above have their limitations. To overcome the limitation, these approaches have been integrated to obtain better performance. Burke [12, 19] presented seven categories of the hybrid recommender system as follows: weighted, switching, mixed, feature combination, cascade, feature augmentation, and model.

Malinowski et al. [17] applied the probabilistic model to two parts of JRS: a CV-recommender and a job recommender separately and integrate the result in order to improve the match between job applicants and jobs. Keim [20] integrated the prior research into a unified multilayer framework supporting the matching of individuals for recruitment and team staffing processes. Fazel-Zarandi and Fox [21] combined different matchmaking strategies in a hybrid approach for matching job applicants and jobs by using logic-based and similarity-based matching.

III. Comparative Study on JRS

The aforementioned recommendation approaches in the JRS are presented for academic research. However, as a practical system, JRS should be analyzed from a product perspective including user profiling, recommendation strategies, recommendation output, and user feedback. A JRS consists of a job applicant subsystem which is designed for job applicants and an e-recruiting subsystem that is used by recruiters. The recommendation principles of two subsystems are basically the same. The scope of this paper lies in the job applicant subsystem owing to a considerable amount of job applicants and its wide range of application in the real world.

Four well known online JRSs, CASPER, Proactive, PROSPECT and eRecruiter, coming from Germany, French, and Hong Kong, are investigated for a comparison purpose. The CASPER is a classical job applicant subsystem that used for enhancing the performance of the JobFinder (http://www.jobfinder.com). The Proactive has different recommendation modules applied to its own website (http://www.proactiverecruitment.co.uk). The PROSPECT is developed by analyzing and mining the resume. The eRecruiter is designed for expanding the functionality and improving the accurateness of the Absolventen.at (http://www.absolventen.at). The comprehensive comparison of four online JRSs is shown in Table I based on their related literatures [14, 22-24] and websites. The usage of four online JRSs' corresponding recruiting websites, which is obtained from Alexa statistics, is shown in Table II. The Proactive and PROSPECT are not in the comparison for the reason that they are lack of the data and the online website, respectively. The XMRC.com (http://www.xmrc.com.cn) is a local e-recruiting website for our case study and details are introduced in Section IV.

TABLE I.
THE COMPARISON OF JOB APPLICANT SUBSYSTEM

Subsection \ System Element		CASPER	Proactive	PROSPECT	eRecruiter
III.A	User Profile	Individual information and behavior	Individual information	Individual information	Individual information and behavior
III.B	Approach	CFR CBR	CBR KBR	CBR	CBR KBR
III.C	Layout	Comprehensive list	Modular list	Comprehensive list	Comprehensive list
III.D	User Behavior	Apply Collect	Apply	Lack of website	Email
Related Literature		Bradley et al. 2003 [25] Rafter et al. 2000 [22]	Lee et al. 2007 [14]	Singh et al. 2010 [23]	Hutterer 2011 [24]

TABLE II.
THE USAGE OF ONLINE RECRUITING WEBSITES

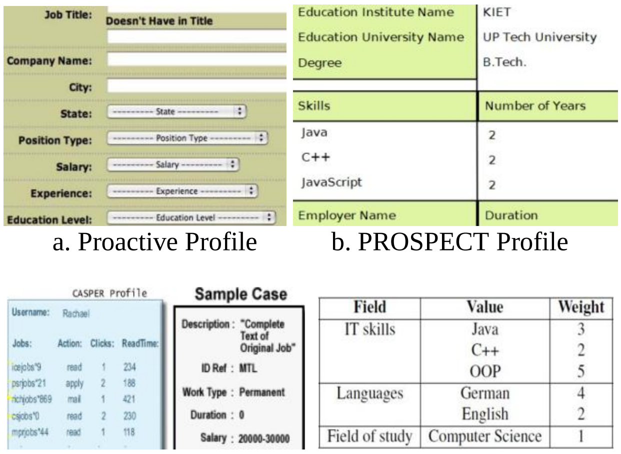
Index \ Website	CASPER (JobFinder)	eRecruiter (absolventen.at)	XMRC.com
Daily IP Visit	1200	3000	63600
Daily Page View	1200	9600	699600
Access Speed	1625Ms/67min	1415Ms/51min	2699Ms/17min
Daily IP Visit	1200	3000	63600

In the following sections, we analyze the differences of four online JRSs from four aspects and summarize their advantages and disadvantages.

A. User Profiling

User profiling is the first step of building a JRS for enhancing the recommendation experience. As the input of the JRS, the user profile captures the main preferences of users and is usually composed of different components. Figure 1 shows that the samples of four online JRSs' user profiles. In terms of the content of the user profile, the Proactive (Figure 1.a) and PROSPECT (Figure 1.b) use the individual information as their user profile, including education experience, working experience and skill. Not only individual information but also historical behaviors such as providing job application and collecting job posts, are considered in the CASPER (Figure 1.c) and eRecruiter (Figure 1.d). More individual information and historical behaviors are collected for profile presentation, both more accurate user preferences and more effective recommendation results will be obtained.

Although all the four online JRSs utilize the individual information, their origins are not the same. The CASPER and Proactive capture the user preference based on the description of a preferred job, while the PROSPECT and eRecruiter mine the resume to generate the user profile. Different from the PROSPECT which mines the resume by using the text mining technology, the eRecruiter represent the resume as a vector model for applying to recommendation algorithm. We can acquire a considerable amount of information about job applicants from different channels.



a. Proactive Profile b. PROSPECT Profile

c. CASPER Profile d. eRecruiter Profile

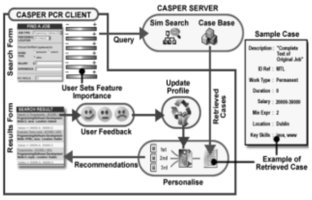
Figure 1. User Profiles of Four Online JRSs

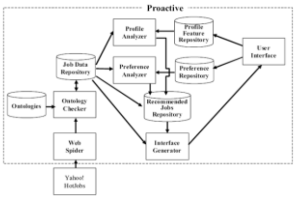
B. Recommendation Strategies

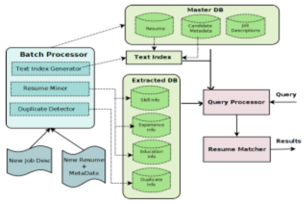
The recommendation strategy refers to the choice of recommendation approaches. Common approaches used in the JRS have been introduced in Section II. Different online JRSs may employ different recommendation approaches based on their own user profilers. The PROSPECT uses a single CBR which has a high requirement on the accuracy of user profiles, while other online JRSs employ two approaches in the form of HyR for recommendation but their categories are not the same. Based on the particular user profile, the CASPER uses the parallel HyR which selects the corresponding approach such as CBR and CFR, respectively. On the contrary, a cascaded HyR which uses KBR and CBR successively is applied in the Proactive and eRecruiter.

Furthermore, the system architecture that describes the information flow and function module of a system can also explain the recommendation strategy. The architectures of four online JRSs are shown in Figure 2.

Figure 2 shows that the information flow of these four online JRSs is common: data collection, data processing, recommendation, and result output. There is another common ground that the recommendation is designed as a module to process the profile and output the result in four online JRSs. In addition, each online JRS has its own additional functions, such as the ontology checker in the Proactive and the resume miner in the PROSPECT.


a. CASPER PCR [22]


b. Proactive [14]


c. PROSPECT [23]

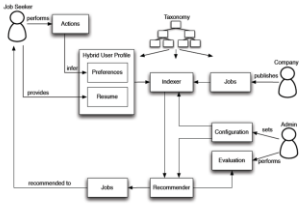

d. eRecruiter [24]

Figure 2. Architectures of Four Online JRSs.

C. Recommendation Output

The recommendation output is usually in the form of a list of jobs, each of which is described briefly in the JRS. It allows the job applicant to have a basic understanding of the recommended job. Besides, the form of “Top-N”, “You Maybe Also Like” and “What Others Looking” are also popular and effective. The four online JRSs use the traditional form to list the recommended job. Their output form is very simple and it is not easy to screen the job that a job applicant is most interested in. Furthermore, recommender explanation (e.g., explaining why the system recommends the jobs) is also an important part of the output but all the existing online JRSs have no attention on it.

D. User Feedback

As a part of the user feedback, the user experience of three online recruiting websites is favorable and their screenshots of the home page are shown in Figure 3, where the PROSPECT is lack of an online recruiting website. On the online recruiting website, some buttons such as apply, collect and email, are designed for every recommended job to record the behavior of job applicants. It is convenient for the job applicant to experience the service provided by the JRS and record the user feedback.

From the above aspects, we analyze four online JRSs on a product level and Table III summarizes briefly their advantages and disadvantages.

After summarizing the pros and cons of online JRSs, we can find out some challenges in building a JRS and distinguish the JRS from the generic RS whose architecture is described in Figure 4. The core of a RS is the recommendation module. After taking the user profile as the input and going through the recommendation module, the RS outputs the recommendation results that satisfy the desires of users. Besides the inner factors, the context affects the operation of the RS and the accuracy of the result. Some specific challenges of the job recommender system are introduced as follows:


CASPER


Proactive


eRecruiter

Figure 3. Screenshots of Online Recruiting Websites.

TABLE III.
ADVANTAGES AND DISADVANTAGES OF ONLINE JRSS

JRS	Advantages	Disadvantages
CASPER	Hybrid profile and approach. User can set the feature importance. Update profile based on user feedback.	Content of profile is simple. Use one way recommendation.
Proactive	Hybrid approach. Provide four recommendation modules. Use ontology to classify jobs.	Single profile. Knowledge engineering problem. Only email about user feedback.
PROSPECT	Resume miner. Batch processing.	Single profile and approach. Simple resume match. Use one way recommendation.
eRecruiter	Hybrid profile and approach. Use ontology to classify jobs and users.	Single method of calculating similarity. Use one way recommendation.

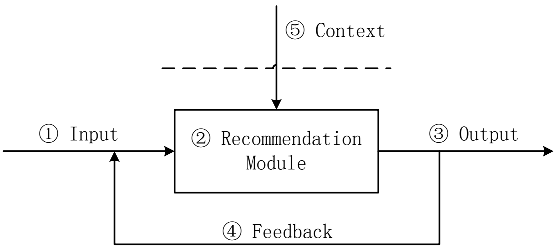


Figure 4. The Typical Architecture of RS.

Timeliness T

The job that posted on the online recruiting website should be timely. A job is unavailable and not recommended to a job applicant when the date has exceeded the recruiting deadline or the number of employees is enough. The timeliness is shown in the input and output module of JRS.

T: (duration,capacity)

The duration is the deadline of a job recruiting, while the capacity is the expected number of employees. The job is available if and only if the day fall into the valid period and the number of employees is less than the value of the capacity.

Reciprocal Recommender

Different from the traditional RS which only considers the unilateral preference, e.g., the preference of a user on the item, the JRS employs a bilateral recommendation approach, which is also called reciprocal recommender. In the JRS, the profile of a job applicant is composed of personal information and job preference while the recruiter's profile consists of self description and job requirement. By integrating the relevance *rel(u~v)* between the job applicant's preference and the recruiter's self description, with the relevance *rel(v~u)* between the job applicant's personal information and the recruiter's requirement, we can obtain the final relevance between *u* and *v* as described as in (1), where *u* is a job applicant and *v* is the recruiter:

$$rel(u,v)=rel(u\sim v)\otimes rel(v\sim u). \qquad (1)$$

Competitiveness C

The competitiveness is defined as the number of job applicants who share the same interest in a job. In the traditional RS, there may be a vast amount of users who have a preference for the same item and the item is recommended to all these users. However, JRS is different from RS due to the competitive relation among job applicants and the timeliness of a job. So the minimal competitive value which is measured by the similarity between the job applicant and a job, as well as the limited number of job applicants who receive a same job, should be considered to prevent the job applicant from making some hopeless attempts in the JRS.

C: (similarity,capacity)

The similarity is the competitiveness of the job applicant for a job and the capacity is the expected number of recommended job applicants. This characteristic is analyzed and applied in the recommendation module.

Rating Cycle

The user rating cycle is the length of time that the job applicant accomplishes all the operations for a job. The operation of a job applicant begins with viewing the job description and ends with commenting, but there are few comments in most cases. Besides, after finding out the preferred job on the recruiting website, the job applicant takes a long time for the offline process, such as resume selection, written examination and interview. As a result, the user rating cycle has the characteristics of long period and few comments in the JRS.

Context

The context is defined as a set of factors of the objective environment, which affects the whole recommendation process including the selection of user profiles, the application of recommendation approaches and the output of recommendation results. For example, one kind of the context is the factor formed in the peak season and the off season. It affects directly the desire of a job applicant. Generally speaking, the generic RS has small influence of the context factor, for instance, the purchase of a book has no obvious peak season and off season. Therefore, the context is one of challenges for adapting the recruiting trend in the JRS.

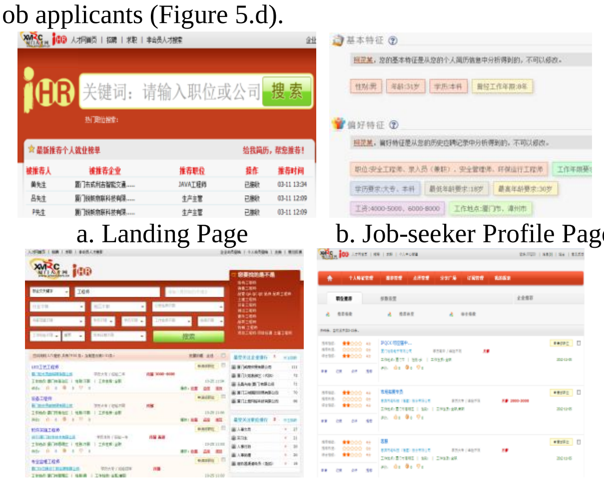
To sum up, the JRS has some specific challenges which are shown in Table IV.

TABLE IV.
THE COMPARISON BETWEEN JRS AND GENERIC RS

	JRS	Generic RS
Timeliness	Yes	No
Reciprocal Recommender	Yes. People-job and enterprise-people.	No. People-item
Competitiveness	Yes	No
Rating Cycle	Long period and few comments.	Short period and many comments
Context	Much	Less

IV. USER CLUSTERING-BASED JRS

As the official website of Xiamen talent service market in China, the XMRC website owns about one and a half millions resumes, while over one hundred and fifty thousand verified job positions are posted every day. However this website also has the challenges described in Table III, it is difficult to design a general recommendation strategy for the JRS. In order to solve this problem, the job applicants are grouped into different clusters where different clusters can use different recommendation approaches. Applying this idea into practical applications, we developed a local JRS called iHR (Figure 5.a) on the platform of the XMRC website. The iHR can extract the user profile automatically (Figure 5.b) and provide the function of searching (Figure 5.c) based on the enormous database. Besides, the iHR provides different lists of recommended jobs for different job applicants (Figure 5.d).



a. Landing Page b. Job-seeker Profile Page

c. Search Page d. Recommendation Result Page

Figure 5. Screenshots of iHR JRS.

In the following, we will force on how to design the user clustering-based JRS on the XMRC website for overcoming the challenges.

A. Problem and definition

Problem: How to classify the job applicants in the JRS and what are the factors affecting the choice of recommendation approaches?

In the iHR, the job applicants are classified into three groups defined as follows:

- *Proactive*: This group of job applicants has a clearly defined goal in finding the job and is active to find out their own preferred jobs by searching or other methods on the recruiting website.
- *Passive*: In this group, the job applicants have no definite ideas about their preferred job. Therefore, they usually turn to other job applicants who may share the same preference for guidance.
- *Moderate*: This group of job applicants is between the "proactive" and "passive". They pay attention to both self-preference and other job applicants' preference on jobs.

We define three features to describe the job applicants which are illustrated as follows:

$$User\ profile = \{U, I, B\}$$

a) User Activity - U

In the JRS, the user activity which indicates the usage of a job applicant for the system is defined by the registration time and the number of logins.

$$U:\ (t, n)$$

Where t is the registration time and n is the number of logins. The user activity is obtained according to these two factors as shown in (2),

$$U\ (t, n) = n/t. \tag{2}$$

b) Information Collection - I

The iHR divides the user individual information into six categories: basic information, educational background, working experience, language skills, job intention, and additional information, in which there are 62 input fields. The information collection is described by the completion of user individual information.

To avoid the deviation caused by the system design, ratio of the number of input fields that completed by the individual user and all users' average number of completed input fields, is used to express the information collection as in (3).

$$I\ (i) = i\ /\ i(A) \tag{3}$$

where i is the number of completed messages and $i(A)$ is the average number of input fields that all the users fulfill in the JRS.

For example, a job applicant fulfills forty input fields of the four categories including basic information, educational background, language skill and job intention, while the average number is fifty. The information collection of this user is $I(i)$=40/50=80%。

c) Behavior Frequency - B

Besides the user individual information, the user behavior as a part of the user profile is also an influencing factor of choosing the recommendation approach. The user behavior consists of clicking, searching and commenting which are recorded in the database. Therefore, we use the click frequency, search frequency and comment frequency to describe the user behavior.

Click Frequency (ClF) The clicking operation describes that the job applicant clicks the buttons of a job on the XMRC website, such as view, apply and collect. As the frequency of clicking operation, ClF is determined by the number of clicking *(c)* and the time *(t1)*.

$$ClF \supset \{\ c, t1\}$$

Search Frequency (SeF) A job applicant can search the preferred job through the search engine in the recruiting website. The searching operation indicates the user preference and nature of a job applicant, as well influence the selection of the recommendation approach. So the number of searching *(s)* in a period *(t2)* is used to define the search frequency.

$$SeF \supset \{\ s, t2\}$$

Comment Frequency (CoF) CoF is obtained by calculating the number of comments *(e)* between job applicants and jobs within the period *(t3)*.

$$CoF \supset \{\ e, t3\}$$

Based on the aforementioned frequency, the behavior frequency is calculated by (4).

$$B(c,s,e)=\{ClF,\ SeF,\ CoF\}=\{c/t1\ , s/t2\ , e/t3\ \}. \tag{4}$$

B. Job Recommendation Based on User Clustering

Based on the defined features of the user profile, we can group the job applicant into three types defined in Section IV(A) by using clustering. They are the proactive group, the moderate group and the passive group. Then we can choose the suitable recommendation approach for each group of job applicants. There are three approaches-CBR, CFR and HyR, which had been achieved in the iHR. The appropriate recommendation approach for a user group is determined according to all of their characteristics. In particular, CBR is suitable for the proactive group, CFR is appropriate for the passive group and the moderate group prefers the result obtained by HyR. This recommendation strategy based on user clustering is different from the traditional strategy which uses a specific approach to recommend items to all types of job applicants. With this new recommendation strategy, a job recommender system becomes more personalized and intelligent.

C. Empirical Evaluation

a) Data set and Normalization

To evaluate the effectiveness of the recommendation strategy of our iHR, we gathered the personnel information of one hundred job applicants ranging from September 1st, 2012 to October 1st, 2012 in the XMRC website database. The personnel information consists of login information, individual information and historical behaviors, and it was used to represent the user profile. The data set were summarized in Table V.

Considering that the unit of multidimensional data had an effect on the data analysis, the data normalization was used to make the data into the common interval for avoiding the dependence of units. The common normalization method is minimum-maximum normalization based on the linear transformation as shown in (5).

$$x(i)'=(\ x(i)-min(f)\ )\ /\ (\ max(f)-min(f)\ ) \tag{5}$$

Where *x(i)* is the original data, *x(i)'* is the normalized data, *min(f)* and *max(f)* are the minimum and maximum of the value of a feature *(f)*, respectively. Table V shows the normalized data of the user profile.

TABLE V.
EXAMPLES OF EXPERIMENTAL DATA

User	User Activity	Information Collection	Click Frequency	Search Frequency	Comment Frequency
Yang	0.6	0.67	0.8	0.6	0.1
Lee	0.8	0.3	0.9	0.7	0.2
Hong	0.2	0.5	0.3	0.4	0.01

b) Experiment

For evaluation purpose, we recorded the login and individual information of one hundred job applicants and gathered their behavior information over a period of a month in the XMRC website. Based on the personnel information and the above defined equations, the user profile of a job applicant which contains five features was calculated for clustering. Grouping one hundred job applicants by k-means, we obtained three groups: the proactive group, the moderate group and the passive group. Afterwards, three recommendation approaches achieved in the iHR were employed in each user cluster for evaluating the recommendation strategy, respectively.

We evaluated the recommendation strategy by measuring the satisfaction rate of the accepted jobs in the list of Top-N recommendation. Given a list of recommendation *R* to a job applicant *U*, the satisfaction rate (shown in (6)) is defined as the proportion of the accepted jobs that the job applicant prefers, where *N(x)* is the number of accepted jobs and *N* is the number of R.

$$Satisfaction\ (X, R) = N(x) / N \qquad (6)$$

In the experiment, we employed CBR, CFR and HyR for recommending jobs (top@5, top@10, top@20 and top@40) to all the job applicants and three user groups, respectively. Their satisfaction rates are shown in Figure 6 and their recall rates are shown in Figure 7. We compared our proposed method (e.g., apply different recommendation approaches for different user groups) with the following baselines:

- applying CBR to all the job applicants,
- applying CFR to all the job applicants,
- applying HyR to all the job applicants.

The comparison of their satisfaction rates are shown in Figure 6.a. From the results, we observed that the satisfaction rate provided by our proposed method is better than the three other baselines. To gain more insights on the choice of recommendation approaches, we compared different user groups employed three different recommendation approaches, such as CBR, CFR and HyR. Our comparative results in Figure 6.b-d indicate that each user cluster has its own appropriate recommendation approach. In particular, Figure 6.b shows that the proactive group is suitable for CBR since the satisfaction rate caused by CBR is higher than the others. Similarly, Figure 6.c and 6.d illustrate that the passive group is appropriate for CFR and the appropriate approach for the moderate group is HyR, respectively.

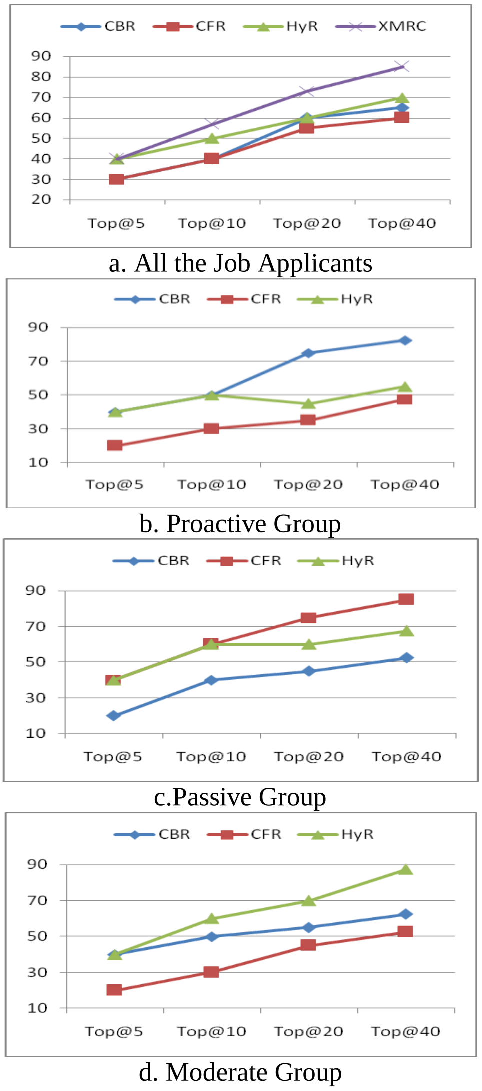


Figure 6. Satisfaction Rates of All the Job Applicants and Three User Groups on Different Recommendation Approaches.

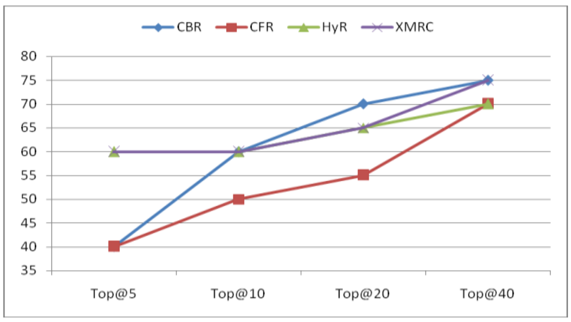


Figure 7. Recall Rates of All the Job Applicants on Different Recommendation Approaches.

V. CONCLUSION AND FUTURE WORK

In this paper, we design, develop and deploy an online JRS for choosing the suitable recommendation approaches based on users' characteristic. To improve the accuracy and effectiveness of our system, we first investigate four existing online JRSs from four different aspects: user profiling, recommendation strategies, recommendation output, and user feedback. We then summarize the advantages and disadvantages of these online JRSs and highlight the differences between the JRS and the generic RS for generalizing the challenges in building high-quality JRSs. To address the challenge caused by a single recommendation approach in a JRS, we group users into different clusters and employ different recommendation approaches for different user clusters.

Besides, the accuracy and effectiveness of the JRS can be largely improved. In particular, the reciprocal recommender can be further applied, e.g. building a bilateral evaluation matrix. We also plan to take the context factor into consideration.

ACKNOWLEDGMENT

The project was supported by the Natural Science Foundation of Fujian Province of China under Grant No. 2011J05157 and by the National Natural Science Foundation of China under Grant No. 61070151.

REFERENCES

[1] S. T. Al-Otaibi and M. Ykhlef, “A survey of job recommender systems,” International Journal of the Physical Sciences, vol. 7(29), pp. 5127-5142, July, 2012.

[2] S. T. Zheng, W. X. Hong, N. Zhang and F. Yang, “Job recommender systems: a survey,” In Proceedings of the 7th International Conference on Computer Science & Education (ICCSE 2012), pp. 920-924, Melbourne, Australia, July, 2012.

[3] M. Gao and Y. Q. Fu, “User-Weight Model for Item-based Recommendation Systems,” Journal of Software, vol. 7(9), pp. 2133-2140, 2012.

[4] K. Yu, G. Guan and M. Zhou, “Resume information extraction with cascaded hybrid model,” In Proceedings of the 43rd Annual Meeting of the ACL, pp. 499-506, Ann Arbor, Michigan, June, 2005.

[5] X. Yi, J. Allan and W. B. Croft, “Matching resumes and jobs based on relevance models,” In Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 809-810, Amsterdam, The Netherlands, 2007.

[6] I. Paparrizos, B. B. Cambazoglu and A. Gionis, “Machine learned job recommendation,” In Proceedings of the fifth ACM Conference on Recommender Systems, pp. 325-328, Chicago, USA, October, 2011.

[7] J. S. Breese, D. Heckerman and C. Kadie, “Empirical analysis of predictive algorithms for collaborative filtering,” In Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence, pp. 42-52, 1998.

[8] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions,” Knowledge and Data Engineering, IEEE Transactions on, vol. 17(6), pp. 734-749, 2005.

[9] H. W. Ye, “A Personalized Collaborative Filtering Recommendation Using Association Rules Mining and Self-Organizing Map,” Journal of Software, vol. 6(4), pp.732-739, 2011.

[10] L. Hu, W. B. Wang, F. Wang, X. L. Zhang and K. Zhao, “The Design and Implementation of Composite Collaborative Filtering Algorithm for Personalized Recommendation,” Journal of Software, vol. 7(9), pp. 2040-2045, 2012.

[11] F. Färber, T. Weitzel and T. Keim, “An automated recommendation approach to selection in personnel recruitment,” In Proceedings of the 2003 Americas Conference on Information Systems, pp. 2329-2339, Tampa, USA, 2003.

[12] R. Burke, “Hybrid recommender systems: survey and experiments,” User Modeling and User-Adapted Interaction, vol. 12(4), pp. 331-370, 2002.

[13] C. F. Chien and L. F. Chen, “Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry,” Expert Systems with Applications, vol. 34(1), pp. 280-290, 2008.

[14] D. H. Lee and P. Brusilovsky, “Fighting information overflow with personalized comprehensive information access: a proactive job recommender,” In Proceedings of the Third International Conference on Autonomic and Autonomous Systems, Washington, DC, USA, 2007.

[15] L. Pizzato, T. Rej, T. Chung, K. Yacef, I. Koprinska and J. Kay, “Reciprocal recommenders,” In Proceedings of 8th Workshop on Intelligent Techniques for Web Personalization and Recommender Systems, held in conjunction with the 18th International Conference on User Modeling, Adaptation and Personalization (UMAP 2010), Hawaii, USA, June, 2010.

[16] H. T. Yu, C. R. Liu and F. Z. Zhang, “Reciprocal recommendation algorithm for the field of recruitment,” Journal of Information & Computational Science, vol. 8(16), pp. 4061-4068, 2011.

[17] J. Malinowski, T. Keim, O. Wendt and T. Weitzel, “Matching people and jobs: a bilateral recommendation approach,” In Proceedings of The 39th Hawaii International Conference on System Sciences, pp. 1-9, Hawaii, USA, 2006.

[18] L. Li and T. Li, “MEET: a generalized framework for reciprocal recommender systems,” In Proceedings of the 21st ACM International Conference on Information and Knowledge Management, pp. 35-44, Hawaii, USA, 2012.

[19] R. Burke, “Hybrid web recommender systems,” The Adaptive Web, vol. 4321, pp. 377-408, 2007.

[20] T. Keim, “Extending the applicability of recommender systems: a multilayer framework for matching human resources,” In Proceedings of 40th Annual Hawaii International Conference on System Sciences, pp. 169-178, January, 2007.

[21] M. Fazel-Zarandi and M. S. Fox, “Semantic matchmaking for job recruitment an ontolgy based hybrid approach,” In Proceedings of the 3rd International Workshop on Service Matchmaking and Resource Retrieval in the Semantic Web at the 8th International Semantic Web Conference, Washington D. C., USA, 2010.

[22] R. Rafter, K. Bradley and B. Smyth, “Personalised retrieval for online recruitment services,” In Proceedings of the 22nd Annual Colloquium on IR Research, Cambridge, UK, 2000.

[23] A. Singh, C. Rose, K. Visweswariah, V. Chenthamarakshan and N. Kambhatla, “PROSPECT: a system for screening candidates for recruitment,” In Proceedings of the 19th ACM International Conference on Information and Knowledge Management, pp. 659-668, Toronto, Canada, 2010.

[24] M. Hutterer, “Enhancing a job recommender with implicit user feedback,” In Fakultät für Informatik, Technischen Universität Wien, 2011.

[25] B. Keith and S. Barry, “Personalized information ordering: a case study in online recruitment,” Knowledge Based Systems, vol. 15(5-6), pp. 269-275, 2003.