Dynamic User Profile-Based Job Recommender System

Wenxing Hong, Siting Zheng, Huan Wang*
School of Information Science and Technology
Xiamen University
Xiamen, China
hwx@xmu.edu.cn

***Abstract*—In this paper, we propose a dynamic user profile-based job recommender system. To address the challenge that the job applicants do not update the user profile in a timely manner, we update and extend the user profile dynamically based on the historical applied jobs and behaviors of job applicants. In particular, the statistical results of basic features in the applied jobs are used to update the job applicants'. In addition, feature selection is employed in the text information of jobs that applied by the job applicant for extending the feature. Then a hybrid recommendation algorithm is employed according to the characteristics of user profiles for achieving the dynamic recommendation.**

***Index Terms*—Dynamic user profile, feature selection, hybrid recommendation, e-recruitment system**

I. INTRODUCTION

Recently, job recommender systems have gained much attention in both academia [1-4] and industries since they select the appropriate items (e.g. jobs or candidates) from a considerable amount of online recruiting information based on the preference of users (e.g. job applicants or recruiters). Applied the technology of personalized recommendation, the job recommender system solves the problem of information overload on the recruiting website. It has the capability of providing the appropriate jobs for the job applicant and recommending some suitable candidates to a recruiter. Therefore, a job recommender system consists of a job applicant subsystem which is designed for job applicants and an e-recruiting subsystem that is used by recruiters. Because the recommendation principles of two subsystems are basically the same, we take the job applicant subsystem which has the widely range of application in the real world for illustration.

The traditional job applicant subsystem takes the personal information and job intention of a job applicant as his/her user profile, and uses them to generate the recommendation result by employing the recommendation algorithms. However, there are some shortcomings in the practical application. One of shortcomings is that the personal information and job intention maybe not true because of the job applicant's cognitive deviation. Besides, the job applicant does not update his/her personal information in general case after entering the information on the recruiting website for the first time. As time goes on, the job intention will be changed as well. So the other shortcoming is that the user profile maybe out of date.

To overcome the aforementioned shortcomings, we employ the dynamic recommendation in a job recommender system in this paper. In particular, we create a dynamic user profile by analyzing the historical information and behaviors of the job applicant. From two perspectives of the time and dimensionality, the user profile which consists of the basic feature and extracted feature is updated and extended.

In summary, the contribution of this paper is threefold:

- Based on the basic features of jobs applied by a job applicant which indicate his/her preference, the basic feature of this job applicant can be updated automatically. Moreover, the basic feature is updated at regular intervals for keeping the information up to date.
- From the perspective of dimensionality, we use the extracted feature for feature selection to extend the number of features. Along with the increasing number of applied jobs, the number of extended features will become greater and they will change.
- According to the characteristics of dynamic user profiles, we employ a hybrid recommendation algorithm, i.e. user-based collaborate filtering algorithm, for improving the accuracy and effectiveness of the recommendation results.

The rest of the paper is organized as follows. Section II presents a literature review about user profiles in the job recommender system. In Section III, we state the problem and some definitions. Section IV describes the dynamic user profile and Section V provides a dynamic recommendation algorithm. Finally, Section VI contains some conclusions plus some ideas for future work.

II. RELATED WORK

The generation of a user profile in the e-recruitment system, i.e. the selection of features of job applicants and job posts, has been researched. Robertson and Smith [5] review the research on major personnel selection methods and attracting job applicants, including applicant perceptions of personnel selection processes. Chien and Chen [6] develop a data mining framework for personnel selection to explore the association rules between personnel characteristics and work behaviors, including work performance and retention.

Besides the structured information is used as the feature, the unstructured information (e.g., resumes) is an important data for the user profile. Yu et al. [7] design a cascaded information extraction framework for mining resumes to support automatic resume management and routing. Singh et al. [8] present PROSECT, a decision support tool that aids in the short listing of candidates for jobs by extracting various pieces of information from the unstructured resumes with the help of statistical data driven techniques.

Different from the traditional static user profile in which the information do not change, Gauch et al [9] create a dynamic user profile with the consideration of the dynamical change of the interests and preferences of a user. The dynamic feature is classified into short-term feature and long-term feature [10-12]. Gonzalez et al. [13] classify the job applicants by using their resume in the method of text mining for improving the resource utilization and demand fulfill in the Resource Planning system that they build. There is an adaptive learning module be used to update the knowledge base dynamically by discovering new information from the resumes in the system.

Besides user profile, the recommendation algorithm is also the core of a recommender system. The hybrid recommendation approach, which approved that has higher accuracy than other recommendation approaches [14, 15], combines the basic approaches, such as content-based and collaborate filtering approach, to overcome the limitations of each approach.

In the e-recruitment domain, the hybrid recommendation approach has been applied widely. Malinowski et al. [16] apply the probabilistic model into two parts of JRS: a CV-recommender and a job recommender separately and integrated the result in order to improve the match between candidates and jobs. Keim [17] integrates the prior research into a unified multilayer framework supporting the matching of individuals for recruitment and team staffing processes. Fazel-Zarandi and Fox [18] combine different matchmaking strategies in a hybrid approach for matching candidates and jobs using logic-based and similarity-based matching.

III. Definitions and Problem

In this section, we define some elements that will be used in next sections and state the problem that discussed in this paper.

Behavior: The e-recruitment system provides some buttons for the job applicant, such as apply, collect and consult. The behavior of applying, collecting or viewing are recorded in the database. And these behaviors indicate that the job applicant has interested in the job, especially application. In verse, the features of applied jobs satisfy the job applicant's desire.

Recommendation List (Job): The recommendation result is shown to the job applicant in the form of list. The job applicant has a brief view on the description of jobs with the basic features including the job name, enterprise, location and work type.

Interested Set (C_1): If a job applicant apply a job, it indicates that the job applicant has much interested in this job, then the job is classified into the interested set.

Uninterested Set (C_2): Except the jobs that the job applicant applies, other jobs are classified into the uninterested set.

PROBLEM: Given a job applicant, his/her user profile is updated and extended dynamically, and then a hybrid recommendation algorithm is employed to generate the results and achieve the dynamic recommendation.

To solve the problem, the solution is divided into two parts: creating the dynamic user profile and achieving the dynamic recommendation. Details on these two parts are introduced in the next sections.

IV. Dynamic User Profile

To create a user profile, all the information and behaviors on the recruiting website which belong to the job applicants could be considered. By analyzing the information, the user profile can be classified into basic feature, extracted feature and behavior feature [Li]. We can update the basic feature and extend the extracted feature to generate the dynamic user profile.

A. Dynamic Basic Feature

The basic features which entered by the job applicant describe his/her personal information and job intention. It can be used as the standard of appropriate jobs by matching the same features of a job post. Table I gives some common basic features of a job applicant and a job post on a recruiting website.

TABLE I. The Basic Features of A Job Applicant and A Job Post

Job Applicant	Job Post
Sex: Male Age: 26 Degree: Master Education: Xiamen University Work_length: 1 year Need_job_type: Full time Need_salary: 8000 Need_location: Xiamen	Job_name: Programmer Salary: 7000 Location: Xiamen Job_type: Full time Need_sex: Unlimited Need_age: 20-35 Need_degree: Master

The basic features of a job applicant have not been updated since the first information entry. But the basic features of the applied jobs reflect the preferences of the job applicant. Therefore, we make a statistic on the basic features of jobs which applied by a job applicant to update his/her basic features. For example, based on the bahaviors of a job applicant, we classify the location and salary of jobs into the interested set and uninterested set, respectively. From Fig. 1, we can get that this job applicant has a favour of Siming, Xiamen, but has no intention to go to Fuzhou, where ○ represents the interested one and × represents the uninterested one. Similarly, we can find the acceptable salary of this job applicant is about 6000 according to the statistics in Fig. 2. Therefore, we can take Siming, Xiamen and 6000 as the basic feature of the job applicant and recommend the jobs that satisfy the desire. In addition, we have a statistic on the degree and work length of the applied jobs in Fig. 3 and Fig. 4. From the results, we can use Bachelor and 2 as the basic feature. The statistics on the job type of applied jobs (shown in Fig. 5) illustrate that the job applicant has much attention to the full-time jobs.

By the same method, we can statistic and analyze other more basic features. From the statistics, the basic features of the job applicant can be updated. However, the statistic results will change with the increasing number of behaviors of a job applicant. So the system makes a statistic at regular intervals, e.g., a week, to generate the dynamic basic features.

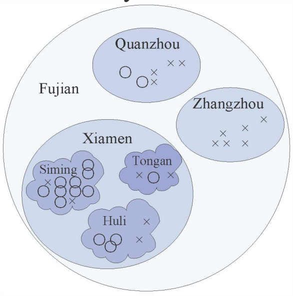

Fig. 1. The statistics of location of historical applied jobs

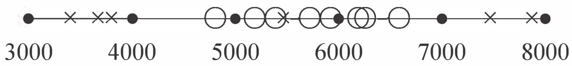

Fig. 2. The statistics of salary of historical applied jobs

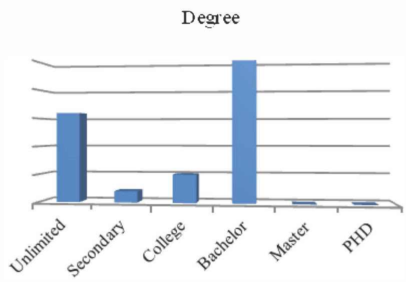

Fig. 3. The statistics of degree of historical applied jobs

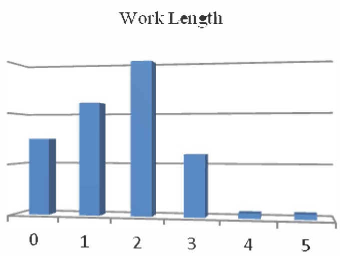

Fig. 4. The statistics of work length of historical applied jobs

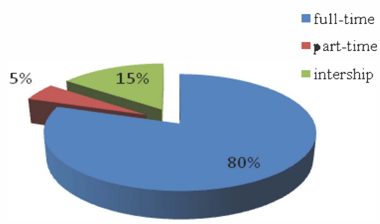

Fig. 5. The statistics of job type of historical applied jobs

B. Dynamic Extracted Feature

The basic features of job posts are used to describe the basic requirements. Besides, the text information including the job description and requirements describe the characteristics of jobs. Therefore, we can extract the feature from the text information to use as extracted features and extend the user profile. With the increasing number of applied jobs, the number of extracted features becomes greater and the dynamic extracted feature is generated from the dimensional perspective. The workflow of dynamic extracted feature is shown in Fig. 6 and introduced as follows:

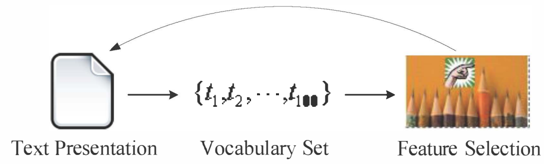

Fig. 6. The workflow of dynamic extracted feature

a) Text Presentation

A text (d) represents the text information of a job in the form of the vector space model after Chinese segmentation and stop words filtration. So the recommendation job set is represented as in (1).

$$Job=\{d_1,d_2,\cdots,d_l\} \tag{1}$$

At the same time, the interested set and uninterested set can be represented by (2) and (3), respectively.

$$C_1=\{d_1^+,d_2^+,\cdots,d_{q_1}^+\} \tag{2}$$

$$C_2=\{d_1^-,d_2^-,\cdots,d_{q_1}^-\} \tag{3}$$

b) Vocabulary Set

Calculate the tf-idf value ω of each word t in the text information of the set Job to use as the weight of the corresponding word. Therefore, the job set is

$$Job=\{t_1,\omega_1,t_2,\omega_2,\cdots,t_l,\omega_l\}. \tag{4}$$

According to the tf-itf value, the top 100 words are selected as the candidate set of extend features and shown in (5).

$$T_c=\{t_1,t_2,\cdots,t_{100}\} \tag{5}$$

c) Feature Selection

Feature selection means selecting the features which have a good capability of distinguishing the category from the candidate set of extend features. And the standard of feature selection is the information gain of each word t in the C_1 and C_2 set.

The core of calculating the information gain is counting the number of information which the feature takes for the classification system. More information is taken, more important the feature is. The formula of calculating the information gain of a word is shown in (6), (7) and (8).

$$GI(t)=H(C)-H(C|t) \tag{6}$$

$$H(C)=-\sum_{i=1}^{2}P(C_i)\times\log_2 P(C_i) \tag{7}$$

Where $P(C_i)$ represents the probability of the document number of the category C_i in all the documents of the classification system.

$$H(C|t)=P_t \times H(C|t)+P_{\bar{t}} \times H(C|\bar{t}) \\ =-P_t \times \sum_{i=1}^{2} P(C_i|t)\times \log_2 P(C_i|t) \\ -P_{\bar{t}} \times \sum_{i=1}^{2} P(C_i|\bar{t})\times \log_2 P(C_i|\bar{t}) \tag{8}$$

Where P_t represents the probability of the documents that contain the feature t. $P(C_i|t)$ represents the probability of the documents of the category C_i that have the feature t in all the documents which contain the same feature t. Similarly, the later half part of the above formula express the probabilities of the documents which have not the feature t.

According to the above formulas, the information gain $GI(t)$ of each word in the candidate set of extended features is calculated. If $GI(t) \geq \theta$, where θ is the predefined index, the word is defined as the extend feature. Finally, the extracted feature set is represented as follow.

$$T=\{t_1,t_2,\cdots,t_n\}$$

V. Dynamic Recommendation

Based on the aforementioned dynamic user profile, we could employ the corresponding recommendation algorithm to achieve the dynamic recommendation for improving the accuracy of recommendation results. Firstly, for solving the cold-start problem, the user-based collaborate filtering algorithm is applied to generate the initial recommendation jobs. So the job applicant can have an operation on the jobs, such as apply and collect. Then the behaviors of job applicants and historical applied jobs are recorded and analyzed to dynamic update and extend the user profile. According to the characteristics of basic features and extracted features in the user profile, we recommend the jobs to the job applicant.

A. Initial Recommendation

Fig. 7 shows the workflow of the initial recommendation which employs the user-based collaborate filtering algorithm and the detail procedure is described as follows:

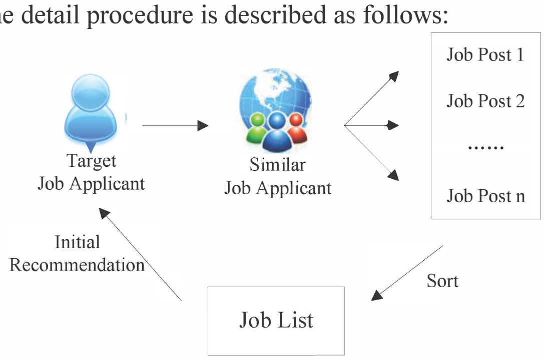

Fig. 7. The workflow of user-based collaborative filtering initial recommendation

Step 1: Calculate the similarity S_i between a target job applicant and other job applicants according to their basic features. Then select the top 100 job applicants based on the similarity as similar job applicants:

$$U_s=\{u_1,u_2,\cdots,u_{100}\}.$$

Step 2: Normalize the similarity of 100 similar job applicants by the normalization equation (9) and the result is used as the weight of each job applicant (u_i).

$$w_i=\frac{S_i\text{-}S_{\min}}{S_{\max}\text{-}S_{\min}},\ i=1,2,\cdots,100 \tag{9}$$

w_i is the job applicant's weight while S_i is the similarity between the target job applicant and the similar job applicant (u_i). $S_{\min}$ and $S_{\max}$ are the minimum and maximum similarity among 100 job applicants.

Step 3: Select the top 5 jobs applicant based on the score of job posts for each similar job. Filter the same jobs and create the initial recommendation job set, $Job=\{J_1,J_2,\cdots,J_n\}$, $n<500$.

Step 4: Calculate the ranking index (δ) in the initial recommendation job set. There are m_l job applicants have the operation on each job (J_l). In reverse, each job has m_l values of similarity between jobs and job applicants, i.e.

$$J_l \rightarrow \{u_{k_1},u_{k_2},\cdots,u_{k_{m_l}}\} \rightarrow \{w_{k_1},w_{k_2},\cdots,w_{k_{m_l}}\}.$$

Where k_r is the number of a job applicant who has an operation on the job (J_l), $r=1,2,\cdots,m_l$. Therefore, the ranking index of each job is shown in (10).

$$\delta_l=\sum^{m_l} w_{k_r},\ r=1,2,\cdots,m_l \tag{10}$$

Step 5: Sort the jobs in the initial recommendation job set according to the ranking index of each job.

B. Dynamic Modification and Extension

After having the initial recommendation results, we can provide them for job applicants and record their behaviors. Then the user profile of the job applicant can be modified and extended by the method that introduced as follows:

Step 1: Get a list of initial recommendation jobs when a job applicant logins. Record the job that the job applicant applies and count the number.

Step 2: The interested set C_1 is generated when the number of applied jobs is more than 10 percent of the set Job. Otherwise, if the number is less than 10 percent of the set, it is need to extend the job to reach the predefined number based on the similar name and basic features of the applied job.

Step 3: The uninterested set C_2 is composed by the jobs which not in the interested set.

Step 4: Analyze the interested set and uninterested set regularly by using the method introduced in Section IV for updating and extending the user profile.

C. Job Recommendation

When the basic features of the job applicant are updated, the new basic features are used to calculate the similarity between the job applicant and the jobs. As a result, the new recommendation results will be more avaiable for the job applicant.

Once the extracted feature is extended, it can be used to calculate the similarity, as well be shown on the user interface as the restraint. The job applicant can use it to filter the jobs for improving the accuracy of the recommendation results.

VI. CONCLUSION AND FURTURE WORK

In this paper, we design a dynamic user profile-based job recommender system. This dynamic job recommender system provides the recommendation jobs that satisfy the changeable preferences of the job applicant. In particular, we update the basic features of a job applicant dynamically based on his/her behaviors and the basic features of applied jobs. Moreover, the extracted features that obtained from the text information of applied jobs are changed and extended dynamically. By combining two features, the dynamic user profile is created and used in the hybrid recommendation algorithm for achieving the dynamic recommendation.

Besides the time and the dimensionality of features, there are other factors affect the dynamic recommendation in the e-recruitment system. For example, the context formed in the peak season and the off season has an influence on the job desire of a job applicant. Therefore, we plan to research and discuss this kind of factors for improving the accuracy of the dynamic recommendation and performance of the e-recruitment system.

ACKNOWLEDGEMENT

The research was supported by the Natural Science Foundation of Fujian Province of China under Grant No. 2011J05157.

REFERENCES

[1] 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.

[2] 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.

[3] 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.

[4] R. Rafter, K. Bradley, and B. Smyth. “Automated collaborative filtering applications for online recruitment services,” Adaptive Hypermedia and Adaptive Web-Based Systems, pp. 363-368, 2000.

[5] I. T. Robertson and M. Smith. “Personnel selection,” Journal of Occupational Organizational Psychology, vol. 74, pp. 441-472, 2001.

[6] 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.

[7] 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.

[8] 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.

[9] S. Gauch, M. Speretta, A. Chandramouli and A. Micarelli. “User profiles for personalized information access,” The Adaptive Web, vol. 4321, pp. 54-89, Springer Berlin Heidelberg, 2007.

[10] H. Kim and P. Chan. “Learning implicit user interest hierarchy for context in personalization,” In Proceedings of IUI’ 03, pp. 101-108, Miami, Florida, January 12-15, 2003.

[11] K. Sugiyama, K. Hatano and M. Yoshikawa. “Adaptive web search based on user profile constructed without any effort from users,” In Proceedings of 13th International Conference on World Wide Web, pp. 675-684, New York, May 17-22, 2004.

[12] D. H. Widyantoro, J. Yin, M. El Nasr, L. Yang, A. Zacchi, and J. Yen. “Alipes: a swift messenger in cyberspace,” In Proceedings of 1999 AAAI Spring Symposium Workshop on Intelligent Agents in Cyberspace, pp. 62-67, Stanford, March 22-24, 1999.

[13] T. Gonzalez, P. Santos, F. Orozco, M. Alcaraz, V. Zaldivar, A. D. Obeso and A. Garcia. “Adaptive employee profile classification for resource planning tool,” In Proceedings of SRII Global Conference 2012 Annual, pp. 544-553, July, 2012.

[14] P. Melville, R. J. Mooney and R. Nagarajan. “Content-boosted collaborative filtering for improved recommendations,” In Proceedings of the Eighteenth National Conference on Artificial Intelligence, pp187-192, Edmonton, Canada, July 2002.

[15] K. Yoshii, M. Goto and K. Komatani. “An efficient hybrid music recommender system using an incrementally trainable probabilistic generative model,” IEEE Transactions on Audio Speech and Language Processing, pp. 435-447, vol. 16(2), 2008.

[16] 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.

[17] 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.

[18] 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.