# A Profile Based Job Recommender System

**Dheeraj Bahl, Vinay Desai, Shreekumar Vibhandik, Isra Fatma**

*Prof. Vijay Kolekar , Dept. of Computer Engineering, K.J. College Of Engg & Mgt. Research, Pune, India*

---

**Abstract -** In today's world there are many job seeking and recruitment websites, we provide a enhanced system which is a hybrid recommender system. The website will be designed for various interactions between the user and system.

It provides organized resources the user needs to express their interests. The system exploits the jobs and user profiles .The system generates personalized recommendations of candidates and jobs. The information collected from the system is modeled using different graphs like directed , weighted, multi-relational graph. The exploiters a 3A ranking algorithm to rank items according to their relevance to the target user.

This data will be then forwarded to the job providers as the recommendations goes.

***Key Words*: Recommendation system; Item-based collaborative filtering; Content-based filtering; Vector Space Model(VSM);Mahout**

## INTRODUCTION

The main idea of our project is to make the recruiters task easy by **creating a automated** job recommendation system based on job-seekers profile. The motivation of our project is to:

The increasing usage of Internet has heightened the need for online job hunting. In the year of 2013, the amount of people who searched jobs on www.ganji.com is almost a billion. According to Jobvites report 2014, 68_ The recommended results can achieve higher score of precision and recall, and they are more relevant with users preferences before.

By creating an easy job recommendation system where everyone can get fair opportunities. The job seeker's profiles will be itered based on their skill set. If the job seeker is a fresher then he will be allotted directly to companies who hires fresher's. The HR will get automated suggestions based on his search of the candidate.

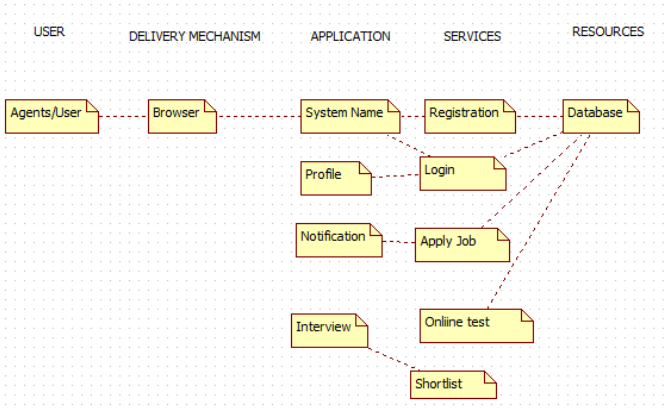


**Fig. 1.1: Basic Job Resume Architecture**

## LITERATURE STUDY

Features of items are abstract and compared with a profile of the users preference. In other words, this algorithm tries to recommend items that are similar to those that a user liked in the past. It is widely applied in information retrieval (IR).

CF is a popular recommendation algorithm that bases its predictions and recommendations on the ratings or behavior of other users in the system. Cosine similarity uses two N-dimensional vectors cosine value to indicate the degree of similarity between them. It is widely used in information retrieval(IR).Tanimoto coefficient, also known as the Jaccard index, measures similarity between finite sample sets, and is denied as the size of the

intersection divided by the size of the union of the sample sets.

**Problem statement:**

To develop the problem under consideration and justify feasibility using concepts of knowledge canvas and IDEA Matrix.

**Feasibility Study:**

Feasibility is the measure that determines the importance of performing the project. The study of the procedure used for determining the feasibility factor is termed as feasibility study. The analyst will consider seven different aspects while conducting the feasibility study given the condition that those aspects are inter-related types of feasibility. This inflicts the fact that this kind of a project can and should be taken.

**Technical Feasibility**

Technical feasibility is a method which makes considerable outputs in a specific time. Technical feasibility is a approach which specifies important equipment and software that the user wants. Here the technical needs of the system may change but has the same value. Response time under specific conditions or deadlines.

The goal of the project is to divide the given data D into pieces that is D1,D2….Dn.

It has to be done in way that the knowledge that is taken from only Di pieces make D easily computable. Undetermined (in the sense that all its possible values are equally likely).. The scheme we are using is called the Threshold scheme. Secret original data is to be constructed by all the participants when k=n.

The essential idea of Adi Shamir's threshold scheme is that 2 points are sufficient to determine a line, 3 points are sufficient to determine a parabola, 4 points to determine a cubic curve and so forth. That is, it takes K points to determine a polynomial of degree K

1 Suppose we want to use a (K; n) threshold scheme to share our secret S , without loss of generality assumed to be an element in a finite set F. Choose random K-1 coefficients a1; :::; ak  1 in F, and leta0=s Build the polynomial

f(x) = a0 + a1X + a2X2 + a3X3 + : + ak 1Xk1:

Let us construct any n points out of it, for instance seti 1;;n to retrieve(i;f (i)) Every participant is given a point (a pair of input to the polynomial and output). Given any subset of k these pairs, we can find the coefficients of the polynomial using interpolation and the secret is the constantterma0.

## SYSTEM ENVIRONMENT

Data Pre-processing: In this step, we clean the raw data to alter useless data including inactive users and expired recruiting information. Construct the item-based CF recommender: In this step, the recommender deals with the input data(apply records) and gives out original predict preference grades using item-based collaborative altering algorithm.

Rescorer: re-compute grades of candidate items. Rescorer: re-compute grades of candidate items.

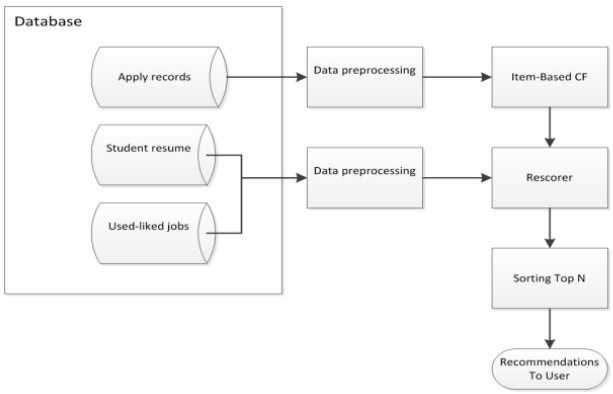


**Fig 3.1 System environment**

## NATURAL LANGUAGE PROCESSING

The traditional Item-based CF processes as follow: First, for each job which current user applied in the past(we regard user-applied jobs as user-liked jobs), find out other users who applied this job(we regard these users as co-applied users), and then and out other jobs these co-applied users also applied, except for the current job(user-liked jobs), uses these jobs as candidate set. The procedure are presented below for each job i user i applied for each co-applied user who applied job find out jobs that user applied; add these jobs to candidate set; delete job i from candidate set.

## CONCLUSION

On the basis of chapter 4, we evaluated user-based CF and item-based CF with different similarity calculation methods. Finally, we selected item-based CF algorithm for its better performance considering various factors. In addition, we also take co-apply users and user apply records in the past into account. The test results indicate the improved recommender with a rescorer is better than a traditional item-based one. Our student job hunting recommender achieved higher precision ,recall and F1 score. Furthermore, the recommended jobs are more relevant with students' preferences. To further optimize the recommendation system and ameliorate the sparsity of user profile, some methods of filling users' preference matrix can be utilized, for example, take advantage of students' implicit behaviors in process of job hunting, which need further research.

## REFERENCES

[1] Schafer J B, Frankowski D, Herlocker J, et al. Collaborative _ltering recommender systems[M]//The adaptive web. Springer Berlin Heidelberg, 2007: 291-324.

[2] Pazzani M J, Billsus D. Content-based recommendation systems[M]//The adaptive web. Springer Berlin Heidelberg, 2007:325-341.

[3] Sarwar B, Karypis G, Konstan J, et al. Item-based collaborative _ltering recommendation algorithms[C]//Proceedings of the 10th international conference on World Wide Web. ACM, 2001: 285-295.

[4] Nikolaos D. Almalis ,Prof. George A. Tsihrintzis , Nikolaos Karagianniset al."FoDRA - A New Content-Based Job Recommendation Algorithm for Job Seeking and Recruiting".

[5] Anika Gupta, Dr. Deepak Garg. "Applying Data Mining Techniques in Job Recommender System for Considering Candidate Job Preferences".

[6] Dunning T. Accurate methods for the statistics of surprise and coincidence[J]. Computational linguistics, 1993, 19(1): 61-74.

[7] Emmanuel Malherbe , Mamadou Diaby , Mario Cataldi et al. "Field Selection for Job Categorization and Recommendation to Social Network Users". 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014).