A hybrid random-walk based web service recommendation enhanced by matrix factorization

Lin Jian
School of Physics and
Electronic Information Engineering,
WenZhou University
Email: neolinjian@gmail.com

Homer Simpson
Twentieth Century Fox
Springfield, USA
Email: homer@thesimpsons.com

Abstract—Recently, the Qos(Qaulity of Serivce) of Web Service that includes response-time, throughput and so on that needs more accuracy prediction. For many web service callers, choosing the appropriate service in right time should be more significant events. So the web serivce recommendation is right to be the choice. The collaborative filtering is major approach to predict the Qos of more web service through the observed data. But the sparse density of data need new technology to enhance the accuracy of prediction. And the matrix factorization is aslo the common measure to solve the prediction. In this paper, we propose the new hybrid approach that combined the predictions with random-walk based and matrix facotrizations. Comprehensive experiments on the QoS data set of real-world web service, that our approach achieve the more accuracy predictions.

Index Terms—random-work, web service recommendation, matrix factorization

I. INTRODUCTION

Overview the past few years, the collaborative filtering and matrix factorization have successed in traditional fields of recommendation, such as Goods, Music, Moive and so on. The recommendation in web service was effected by the achievements. However, the scenario in web service is more complex that suffers from sparse data and incomplete related information. There are so many different web services distributting over heterogeneous network which contains several auto-systems. So the recommendation in web service should solve the problems that sparse QoS(Quality of Service) value collected from various with the untrusted infomation about location or network. In a word, more measures should be made to enhanced the limited information to achieve the more accuracy rate of web service recommendation. Only that, the system of web service can provide the more quality service.

Web Service QoS predicted information enhanced technology is developing fastly. For example, time-aware recommendation that makes prediction by history call record, location-aware recommendation that make use of numbers of AS(auto system), IP or GPS(Global Position System). But the measures all achieve improvement in accuracy rate of prediction in small scale with the sparse data. Athough the information is critical to prediction, the experiments prove the factor that the more appropriate neighhorhood ranking can really boost the accuracy rate of prediction. So the paper that Random Walk Models can efficiently work in real-world datasets in the past

year. With the transition probability matrix which based on the principle of markov random process, the undirected connected users can calcuate the similarity for neighborhood selection.

In the field of web service recommendation, the random-walk models is efficient, but the accuracy rate of prediction need more improvement. The matrix factorization had ever solved the sparse efficiently in similar scenario. Naturally, we will try to combine the random-walk model with the matrix factorization with the good performance. And the matrix factorization also is the best measure to reduction of dimensions, when we calculate the similarity between user and user, the time complexity will be smaller. With high-efficiency algorithm, the hybrid algorithm improve the accuracy rate of prediction in final.

In summary, to solve the web service recommendation and to increase the accuracy rate of Qos prediction, in this paper, the contributions we made as following:

- We find the latent dimension of matrix factorization with statistic method, and explore the extreme rate of data mining in known probability.
- We propose the hybird approach to combine the user-based collaborative filtering with matrix factorization.
- We conduct the experiments on real-world datasets, and achieve the best accuracy rate of QoS prediction.

The rest of this paper is organized as follows. Section II summarizes the related work and our thought about sparse dataset. Section III introduces our approach to combine the CF and MF algorithm. Section IV reports the experiments and analyst and compare the result of approaches. Section V concludes the paper and discusses the future work.

II. RELATED WORK

In this section, we will introduce the initition of spares density data, and explore the extreme rate of data mining in ideal environment that the sampling rate given in advance. Then the review of technology of recommendations will be displayed, includes collborative filting, matrix factorization, and random-walk model.

A. Initition of sparse density data

In the real-world dataset environment, our recommendation system samples the whole dateset with d density. Suppose that

the Matrix have m users, n services, and the $Q \in \mathbf{R}^{m \times n}$. q_{ij} means the qos of user i called service j.

The relationship between the method choosen and accuracy rate of prediction is clearly displayed.

B. The extreme rate of data mining

Subsection text here.

C. User-based Collaborative Filtering

The CF(Collaborative Filtering)-based have been widely used.

D. Matrix facotorization

The MF(Matrix Facotrization) has also been chosen for its accuracy.

E. Random-Walk model

The random-walk that based on RankPage alike measures (the hidden Markov chain theory) to get more appropriate neighbors ranking with the transition matrix.

III. HYBRID APPOACH WITH RW AND MF

Ssdfsdfsdfsdfsdf

IV. EXPERIMENT AND EVALUATION

A. Dataset and Description

Subsection text here.

B. Evaluation Metric and Parameter

Subsection text here.

C. Result Accuracy rate and Comparision

Subsection text here.

D. Analysis and Deduction

Subsection text here.

V. CONCLUSION

The conclusion can be summarized by this.

ACKNOWLEDGMENT

The authors would like to thank...

REFERENCES

[1] H. Kopka and P. W. Daly, *A Guide to LaTeX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.