Poster: Group Preference based API Recommendation via Heterogeneous Information Network

Fenfang Xie, Liang Chen, Dongding Lin, Chuan Chen, Zibin Zheng, Xiaola Lin
School of Data and Computer Science, National Engineering Research Center of Digital Life
Sun Yat-sen University, Guangzhou, China
{xieff5,lindd5}@mail2.sysu.edu.cn,{chenliang6,chenchuan,zhzibin,linxl}@mail.sysu.edu.cn

ABSTRACT

Heterogeneous information networks (HINs) are logical networks which involve multiple types of objects and multiple types of links denoting different relations. Previous API recommendation studies mainly focus on homogeneous networks or few kinds of relations rather than exploiting the rich heterogeneous information. In this paper, we propose a mashup group preference based API recommendation method for mashup creation. Based on the historical invocation experience, different semantic meanings behind meta paths, hybrid similarity measurement and the rich interactions among mashups, we build the API recommendation model and employ the model to make personalized API recommendation for different mashup developers. Extensive experimental results validate the effectiveness of our proposed approach in terms of different kinds of evaluation metrics.

KEYWORDS

Heterogeneous information network, meta path, API recommendation, group preference

ACM Reference Format:
Fenfang Xie, Liang Chen, Dongding Lin, Chuan Chen, Zibin Zheng, Xiaola Lin. 2018. Poster: Group Preference based API Recommendation via Heterogeneous Information Network. In *ICSE '18 Companion: 40th International Conference on Software Engineering Companion, May 27-June 3, 2018, Gothenburg, Sweden,* Jennifer B. Sartor, Theo D'Hondt, and Wolfgang De Meuter (Eds.). ACM, New York, NY, USA, Article 4, 2 pages. https://doi.org/10.1145/3183440.3195010

1 INTRODUCTION

Mashups are an exciting genre of interactive Web applications that draw upon content retrieved from external data sources to create entirely new and innovative services [3] . Mashups can accelerate development period and enhance scalability. The exponential growth of the number of APIs results in unprecedentedly large scope of choices on selecting APIs. Thus, it has become more difficult than ever to build innovative mashup applications. In order to shorten the mashup development period, it becomes a significant challenge to effectively recommend mashup developers with appropriate APIs from a large number of APIs.

ICSE '18 Companion, May 27-June 3, 2018, Gothenburg, Sweden

ACM ISBN 978-1-4503-5663-3/18/05.
https://doi.org/10.1145/3183440.3195010

Recently, most of research works in API recommendation have been focused on semantic based API recommendation [4], Quality of Services (QoS) based API recommendation [1, 7] and social network based API recommendation [6]. However, the shortcoming of the semantic based API recommendation approaches is the singleness of employed data. The weakness of QoS based API recommendation approaches is the instability of QoS information and the difficulty of collecting QoS value of APIs. The approaches based on social networks can not always be applied to API recommendation in a very appropriate manner due to the limited application scenario. Currently, most studies have diverted their attention into recommending APIs via employing the information networks [2]. However, most of information networks based approaches only take advantage of a few related attributes of mashups/APIs and relationships between mashups/APIs, and lack considering the interactions among mashups and APIs. Which highly limits the effect of these information networks based approaches.

To help mashup developers find appropriate APIs, in this paper, we propose a mashup Group Preference based API Recommendation method for mashup creation, called GPRec. GPRec utilizes the historical mashup-API invocation experience, different semantic meaning of meta paths, hybrid similarity measurement and the rich interactions among mashups to build the API recommendation model. And then, the GPRec model is employed to make personalized API recommendation for different mashup developers.

2 APPROACH

As can be seen in Figure 1, our proposed approach consists of three parts. The details of each part are as follows:

(1) **Input Information**. Mashup information (e.g., tag, category, description), API information (e.g., tag, category, description, provider) and historical invocation records between mashups and APIs are regraded as input parameters for GPRec model.
(2) **GPRec Model**. After a series of input handling, mashups, APIs and their related attributes form a heterogeneous information network. Mashups can be connected through multiple typed meta paths representing diverse semantic meaning. Next, a hybrid similarity measurement is considered to compute the similarity between mashups by integrating four different similarity measurement approaches. Finally, a group preference Baysesian personalized ranking algorithm [5] is applied to rank every pair of mashup-API.
(3) **Output Information**. After learning the model by mashup group preference Baysesian personalized ranking algorithm, a list of personalized ranking results between mashups and APIs are obtained.

Table 1: Recommendation Performance Comparison

Top-N	Metrics	BPRKNN	SVD	BPRSVD	POPRank	PaSRec	**GPRec**
Top-3	Precision	0.0587	0.1448	0.1519	0.1566	0.3405	**0.3476**
	MRR	0.0820	0.2697	0.2811	0.3198	0.7240	**0.7332**
	DCG	0.1135	0.3751	0.3952	0.4238	1.0695	**1.0888**
Top-5	Precision	0.0529	0.1348	0.1369	0.1350	0.2529	**0.2593**
	MRR	0.1022	0.3230	0.3342	0.3675	0.7792	**0.7910**
	DCG	0.1657	0.5623	0.5835	0.5932	1.3554	**1.3919**
Top-10	Precision	0.0379	0.0902	0.0930	0.0928	0.1591	**0.1652**
	MRR	0.1179	0.3533	0.3681	0.4029	0.8230	**0.8392**
	DCG	0.2382	0.7257	0.7625	0.7846	1.7437	**1.8288**

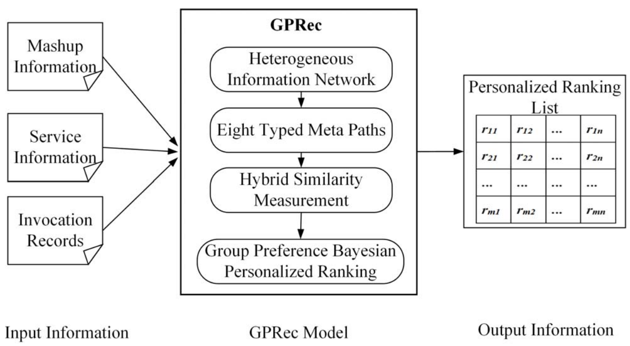

Figure 1: Architecture of GPRec

3 EXPERIMENT

In our experiments, the real-world dataset is crawled from ProgrammableWeb (PW). It contains a collection of 6958 mashups, 1271 APIs and 14731 interactions between mashups and APIs. The detailed statistics of the dataset are shown in Table 2.

Table 2: Statistics of Our PW Dataset

Statistics	Mashup	API
Number of Entities	6958	1271
Number of Categories	373	60
Number of Tags	2089	975
Number of Providers	0	970

To observe performance comparison results between all approaches intuitively, the results under the training data density of 80 percent are shown in Table 1. As can be seen in Table 1, our GPRec model outperforms the other five state-of-the-art approaches in terms of three metrics (e.g., Precision, MRR and DCG). Specifically, BPRKNN performs the worst among all the baselines due to lack of considering other useful information (e.g., network information, description information). POPRank, SVD and BPRSVD obtain worse performance, wherein, BPRSVD is better than SVD because BPRSVD introduces the idea of pairwise based ranking and applys BPR learning algorithm for implicit feedback. POPRank is better than BPRSVD because mashup developers incline to invoke popular APIs. The performance of PaSRec is better than all of the other four methods due to exploiting the heterogeneous information of mashups, APIs and their related attributes. Our GPRec model achieves better performance than PaSRec due to considering more attributes of mashups and APIs, measuring the similarity between mashups from different aspects, and exploiting the mashup group preference.

4 CONCLUSIONS

In this paper, we propose an API recommendation method for mashup creation with mashup group preference. We firstly consider the historical invocation experience, mashups, APIs and their related attributes. Then, We employ a hybrid similarity measurement to measure the similarity between two mashups. Finally, We use a mashup group preference Baysesian personalized ranking algorithm to learn the model and make personalized API recommendation. Comprehensive experiments are conducted on a real-world dataset, which demonstrate the effectiveness of our GPRec model.

ACKNOWLEDGMENTS

The work described in this paper was supported by the National Key Research and Development Program (2016YFB1000101), and the National Natural Science Foundation of China (61722214, 61702568). Liang Chen and Zibin Zheng are the corresponding authors.

REFERENCES

[1] Pinjia He, Jieming Zhu, Zibin Zheng, Jianlong Xu, and Michael R Lyu. 2014. Location-based Herarchical Matrix Factorization for Web Service Recommendation. In *2014 IEEE International Conference on Web Services (ICWS)*. IEEE, 297–304.

[2] Tingting Liang, Liang Chen, Jian Wu, Hai Dong, and Athman Bouguettaya. 2016. Meta-path based Service Recommendation in Heterogeneous Information Networks. In *International Conference on Service-Oriented Computing*. Springer, 371–386.

[3] Duane Merrill. 2006. Mashups: The new breed of Web app. *IBM Web Architecture Technical Library* (2006), 1–13.

[4] Hafida Naïm, Mustapha Aznag, Nicolas Durand, and Mohamed Quafafou. 2016. Semantic Pattern Mining Based Web Service Recommendation. In *International Conference on Service-Oriented Computing*. Springer, 417–432.

[5] Weike Pan and Li Chen. 2013. GBPR: Group Preference Based Bayesian Personalized Ranking for One-Class Collaborative Filtering. In *International Joint Conferences on Artificial Intelligence*, Vol. 13. 2691–2697.

[6] Yan Wang, Lei Li, and Guanfeng Liu. 2015. Social Context-Aware Trust Inference for Trust Enhancement in Social Network based Recommendations on Service Providers. *World Wide Web* 18, 1 (2015), 159–184.

[7] Zibin Zheng, Hao Ma, Michael R Lyu, and Irwin King. 2011. Qos-Aware Web Service Recommendation by Collaborative Filtering. *IEEE Transactions on services computing* 4, 2 (2011), 140–152.