

A Web service recommendation system based on QoS attribute based collaborative filtering

S. Aruna

Department of CSE, Mepco Schlenk Engineering College, Sivakasi, Tamilnadu.

E-mail: aruna@mepcoeng.ac.in

P. Abinaya

Department of CSE, Mepco Schlenk Engineering College, Sivakasi, Tamilnadu.

E-mail: abinayap@mepcoeng.ac.in

Received January 12, 2020; Accepted February 14, 2020

ISSN: 1735-188X

DOI: 10.14704/WEB/V17I1/WEB17001

Abstract

Service recommendation systems are an effective method for predicting user preference for services. Of late collaborative filtering based methods are popularly used for service recommendation. A similarity computation based on QoS is used, where attribute values between users or items is compared. The service recommendation system is also improvised by utilizing the QoS attributes and response time to predict the unknown value.

Keywords

QoS, Web service, SOA

Introduction

WEB service is a software system designed to support interoperable machine-to-machine interaction over a network. With the prevalence of Service-Oriented Architecture (SOA), more and more Internet applications are constructed by composing Web services. As a consequence, number of Web services has increased rapidly over the last decade. Web service discovery has become a crucial and challenging task for users. With the large increase in the number of users and a rapid increase in the number of services, it is complicated and time consuming, ineffective to identify the best optimal service. Based on personal preference of user and past preference records it is an attractive solution to recommend unused suitable services.

There are different methods for service recommendation. They are collaborative filtering, content based filtering and Hybrid filtering technique. Among these collaborative filtering (CF)-based service recommendation system is the most widely used. A CF based recommending system need not examine the content of the item. It depends on the user's evaluation of an item and recommends item or a service to similar set of users. Collaborative filtering is based on an assumption that if a group of users liked same things they might concurrently like other items. For example a group of users watched a movie and provided rating(similar evaluations)

Service selection so far has not considered non functional requirements. An optimal service recommendation solution should however consider both functional and non-functional attributes of services, that is (QoS)

Methodologies

Existing system

Service recommendation systems are of several types such as context based, trust based, relationship based, collaborative filter based. Among these the most popularly used is collaborative filter based recommendation system. Pearson correlation coefficient(PCC) and cosine(COS) methods are the commonly used methods to calculate similarity. These methods suffer from limited accuracy. The traditional methods of similarity computation gave less consideration to QoS. Experimental results for PCC and COS methods to provide recommendations were not without fault. The exiting methodologies did not accurate recommendations for unknown value. Also the previous QoS prediction methods assume that the co-invoked Web services have equal contribution weights when computing similarity between two users. However the personalized characteristics (e.g., QoS variation) of both Web services and users should be incorporated into measuring the similarity among users and services.

The similarity between two users u and v is calculated using PCC as below:

$$sim(u, v) = \frac{\sum_{i \in I} (r_{u,i} - \bar{r}_u)(r_{v,i} - \bar{r}_v)}{\sqrt{\sum_{i \in I} (r_{u,i} - \bar{r}_u)^2 (r_{v,i} - \bar{r}_v)^2}}$$

$r_{u,i}$ is the value of item i invoked by user u and \bar{r}_u denotes the average value when user u invoked all of the items in I

The same way PCC can be applied in an item-based collaborative filtering algorithm. The similarity calculation between items i and j is given below as:

$$sim(i, j) = \frac{\sum_{i \in I} (r_{u,i} - \bar{r}_i)(r_{v,i} - \bar{r}_j)}{\sqrt{\sum_{i \in I} (r_{u,i} - \bar{r}_i)^2 (r_{v,i} - \bar{r}_j)^2}}$$

From the above formula $r_{u,i}$ denotes the value when user u invoked item i and \bar{r}_i is the average value when all of the users in u invoked item i . The problem with this approach is that in some cases two users or items may not be similar but might have close values in user-item matrix.. In that case the similarity value computation using PCC method might give invalid values.

To eradicate the inaccurate predictions using the traditional methods the QOS based prediction was proposed and ratio based similarity computation is used.

Proposed System

In order to improve the accuracy and efficiency of the recommendation a new method is used. First a ratio based method is used to compute similarity. The similarity is computed directly by comparing the attribute values of users or items directly. Based on the similarity computation in the first method, new method is used to compute the unknown values.

The proposed system is implemented using collaborative filtering algorithm and experimental result are compared with traditional system. The algorithm is based on prediction and provides recommendations on the ratings or behavior of other users opinion and aggregating them to provide reasonable prediction of active user's preference.

System Flow Diagram

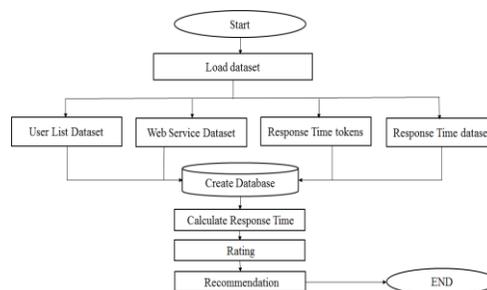


Figure1 System flow Diagram

The initial step is to load the web services dataset. The dataset consist of namely user list dataset and web services dataset. The user list dataset contains user id, ratings given by user for the items or services used. Based on the ratings provided bythe existing users for different web services a database is constructed. Similarity is computed between users and services based on the rating provided by users. Upon computing the similarity unknown values are predicted and services are recommended for users. Finally response time is computed for predicting the recommended services for a new user.

Dataset Description

The proposed methodology is evaluated using real world dataset called the web services dataset. Web services are a budding in recent times and acts as the core technique for building distributed systems. Information about web services are found in Universal Description, Discovery and Integration (UDDI). This aids companies to publish their services. The system uses two types of dataset namely user list dataset and web service dataset derived from XML files in UDDI. User list dataset contains user id, IP Address, Country. The latter consist of Service ID, Website sdl address, Service provider, country etc.

Similarity computation

QoS attribute factors experienced by the user are given in the form of non negative numerical values Similarity is calculated between two users who invoke the same item. This represents the degree of consistency between objects. Consistency can be evaluated using ratio between two values as rated by the user. Adding up all similarities and getting the average final similarity can be calculated.

The method to compute similarity is given as below:

$$Sim(u,v) = \frac{\sum_{i \in I} \frac{\min(r_{u,i}, r_{v,i})}{\max(r_{u,i}, r_{v,i})}}{|I|}$$

Here $r_{u,i}$ and $r_{v,i}$ are the values when users u and v invoked item i . $\min(r_{u,i}, r_{v,i})$ is the minimum of $r_{u,i}$ and $r_{v,i}$ and $\max(r_{u,i}, r_{v,i})$ is the maximum of them.

Likewise the similarity between two items I and j is given by:

$$Sim(i,j) = \frac{\sum_{u \in U} \frac{\min(r_{u,i}, r_{u,j})}{\max(r_{u,i}, r_{u,j})}}{|U|}$$

$r_{u,i}$ and $r_{u,j}$ represent the values obtained by user u when he invoked item i and item j , respectively.

The values $sim(u,v)$ and $sim(i,j)$ are always in the interval of $[0,1]$. The similarity value is computed using the formulae above and the results are judged against other existing methods such as PCC and COS. The similarity computation is applicable to QoS attribute values expressed as numerical values. Despite the QoS values that are given as non-numerical values, such as “excellent” and “Very good”. These subjective ratings are transformed into numerical values according to certain rules and then the proposed similarity computation method can be applied. The types of QoS attributes are classified as below:

Table 1 Three types of QoS attributes

Dependency	Attribute classification	Examples
Do not depend on the user	First type	Throughput, price, Service location
Objectively depend on the user	Second type	Delay, Response time
Subjectively depend on the user	Third type	Degree of user satisfaction

Unknown value prediction

Once similarity is computed between users or items, this acts as basis for predicting unknown values for QoS attributes. The goal of service recommendation system is not only recommending user with appropriate service like other similar users have experienced, but also recommending services that the user has not experienced. To recommend unused service to users it is necessary to predict the unknown values. After summarizing the values the services are recommended. The prediction of unknown value is based on user-item matrix, that contains the response time value of M users for N services. For a current user who has not invoked a service i , the aim is to predict the value of u for an item/service i . the values are predicted using MEAN method where the average

of all services invoked by user u other than service i is calculated. The mean value is taken as the prediction result.

Response time

Response time is the total amount of time it takes to respond to a request for service. The length of time taken for a person or system to react to a given stimulus or event. The response time is calculated based on the user id and service id. The response time is then calculated for proper value of web service address.

Results and Discussion

An efficient service recommendation system is implemented using personalized location aware collaborative filtering system. The highlight of the work lies in recommending services based on non negative QoS attribute values. Computing similarity and prediction for unknown value are the two important procedures in a memory based collaborative filtering service recommendation system. The system takes web services dataset as input and provides recommendation based on similarity values. A new method was utilized to compute similarity. The system takes into account the personal QoS characteristics of both Web services and users to compute similarity between them. This method also incorporate the locations of both Web services and users into similar neighbor selection, for both Web services and users. Comprehensive experiments conducted on a real Web service dataset indicate that the new method significantly outperforms previous CF-based Web service recommendation methods. The following figures represent the dataset and the working of the proposed system.

USER ID	IP ADD...	COUNT...	CONTI...	AS	LATITU...	LONGTI...	REGION	CITY
0	12.108...	United ...	North A...	AS7018...	40.3908	-74.1116	New Jer...	Middlet...
1	12.46.1...	United ...	North A...	AS7018...	40.3908	-74.1116	New Jer...	Middlet...
2	122.1.1...	Japan	Asia	AS4713...	35.69	139.75	Tokyo-to	Tokyo
3	128.10...	United ...	North A...	AS17 P...	40.4167	-86.8751	Indiana	West L...
4	128.10...	United ...	North A...	AS17 P...	40.4167	-86.8751	Indiana	West L...
5	128.11...	United ...	North A...	AS131 ...	34.4228	-119.70...	California	Santa B...
6	128.11...	United ...	North A...	AS131 ...	34.4228	-119.70...	California	Santa B...
7	128.11...	United ...	North A...	AS131 ...	34.4228	-119.70...	California	Santa B...
8	128.11...	United ...	North A...	AS88 Pr...	40.3679	-74.6543	New Jer...	Princeton
9	128.11...	United ...	North A...	AS88 Pr...	40.3679	-74.6543	New Jer...	Princeton
10	128.11...	United ...	North A...	AS5739...	37.0412	-122.124	California	Santa C...
11	128.11...	United ...	North A...	AS5739...	37.0412	-122.124	California	Santa C...
12	128.11...	United ...	North A...	AS5739...	37.0412	-122.124	California	Santa C...
13	128.11...	United ...	North A...	AS1249...	42.3773	-72.4671	Massac...	Amherst
14	128.11...	United ...	North A...	AS1249...	42.3773	-72.4671	Massac...	Amherst
15	128.13...	United ...	North A...	AS160	41.9022	-87.6726	Illinois	Chicano

Figure 2 Sample userlist Dataset

SERV...	WSD...	SERV...	IP AD...	COU...	CONT...	AS	LATIT...	LON...	REGL...	CITY
0	http://...	no-ip...	8.23.2...	Unite...	North...	AS33...	39.89...	-105...	Color...	Broo...
1	http://...	no-ip...	8.23.2...	Unite...	North...	AS33...	39.89...	-105...	Color...	Broo...
2	http://...	no-ip...	8.23.2...	Unite...	North...	AS33...	39.89...	-105...	Color...	Broo...
3	http://...	no-ip...	8.23.2...	Unite...	North...	AS33...	39.89...	-105...	Color...	Broo...
4	http://...	no-ip...	8.23.2...	Unite...	North...	AS33...	39.89...	-105...	Color...	Broo...
5	http://...	dyn dn...	204.1...	Unite...	North...	AS33...	40.79...	-74.0...	New J...	North...
6	http://...	utn.ed...	null	null	null	null	null	null	null	null
7	http://...	utn.ed...	null	null	null	null	null	null	null	null
8	http://...	200.4...	200.4...	Argen...	South...	AS73...	-34.61	-58.3	Distr...	Buen...
9	http://x...	hoteld...	200.5...	Argen...	South...	AS11...	-34.61	-58.3	Distr...	Buen...
10	http://...	libros...	null	null	null	null	null	null	null	null
11	http://...	mob il...	null	null	null	null	null	null	null	null
12	http://...	mstat...	65.55...	Unite...	North...	AS80...	47.67...	-122...	Wash...	Red...
13	http://...	neodv...	190.2...	Argen...	South...	AS16...	-34.61	-58.3	Distr...	Buen...

Figure 3 Sample Webservices dataset

LONGITUDE	REGION	CITY	WSDL ADD...	SERVICE P...	RESPONSE ...	RATINGS
-74.1116	New Jersey	Middletown	http://ewave....	no-ip.com	5.982	1
-74.1116	New Jersey	Middletown	http://ewave....	no-ip.com	0.228	10
-74.1116	New Jersey	Middletown	http://ewave....	no-ip.com	0.237	10
-74.1116	New Jersey	Middletown	http://ewave....	no-ip.com	0.221	10
-74.1116	New Jersey	Middletown	http://ewave....	no-ip.com	0.222	10
-74.1116	New Jersey	Middletown	http://aquest...	dyn dns.org	0.527	9
-74.1116	New Jersey	Middletown	http://www.ut...	utn.edu.ar	0.453	9
-74.1116	New Jersey	Middletown	http://www.ut...	utn.edu.ar	0.566	9
-74.1116	New Jersey	Middletown	http://200.45...	200.45.113...	0.386	10

Figure 4 Collaborative based similarity computation

USER ID	SERVICE ID	RESPONSE TIME	CITY	WSDL ADDRESS	RANKING
12	22	0.171	Santa Cruz	http://www.indu...	10
12	23	0.173	Santa Cruz	http://www.indu...	10
12	25	0.174	Santa Cruz	http://www.micr...	10
12	26	0.174	Santa Cruz	http://www.micr...	10
12	27	0.174	Santa Cruz	http://www.micr...	10
12	19	0.184	Santa Cruz	http://www.enter...	10
12	47	0.197	Santa Cruz	http://api.godo.c...	10
12	49	0.199	Santa Cruz	http://www.mp3...	10
12	35	0.219	Santa Cruz	http://www.stpat...	10
12	20	0.233	Santa Cruz	http://www.enter...	10
12	46	0.235	Santa Cruz	http://api.godo.c...	10
12	29	0.24	Santa Cruz	http://www.touri...	10
12	30	0.243	Santa Cruz	http://www.touri...	10
12	45	0.256	Santa Cruz	http://api.godo.c...	10
12	3	0.262	Santa Cruz	http://ewave.no-i...	10
12	4	0.263	Santa Cruz	http://ewave.no-i...	10

Figure 5 Web service recommendation for different users based on similarity computation

USER ID	SERVICE ID	RESPONSE TIME	CITY	WSDL ADDRESS	RANKING
15	4	0.203	Chicago	http://ewave.no-i...	10
15	3	0.205	Chicago	http://ewave.no-i...	10
15	1	0.212	Chicago	http://ewave.no-i...	10
15	27	0.218	Chicago	http://www.micr...	10
15	26	0.219	Chicago	http://www.micr...	10
15	25	0.221	Chicago	http://www.micr...	10
15	49	0.221	Chicago	http://www.mp3...	10
15	22	0.224	Chicago	http://www.indu...	10
15	23	0.224	Chicago	http://www.indu...	10
15	20	0.232	Chicago	http://www.enter...	10
15	19	0.233	Chicago	http://www.enter...	10
15	2	0.236	Chicago	http://ewave.no-i...	10
15	45	0.252	Chicago	http://api.godo.c...	10
15	47	0.256	Chicago	http://api.godo.c...	10
15	46	0.259	Chicago	http://api.godo.c...	10
15	35	0.272	Chicago	http://www.stpat...	10

Figure 6 Webservice recommendations for another user

USER ID	SERVICE ID	RESPONSE TIME
0	12	7.067
0	13	6.415
0	14	5.828
0	15	5.875
0	16	5.782
0	17	5.776
0	18	5.565
0	19	0.239
0	20	0.24
0	21	6.016
0	22	0.247
0	23	0.247
0	24	5.611
0	25	0.25
0	26	0.249
0	27	0.249
-	--	---

Figure 7 Response time calculations for service recommendation

Similarity was computed based on QoS attributes. Experimental results prove that QoS attribute value similarity computation outperforms other traditional methods in terms of response time and predicting services for unknown services. Existing method is compared against the new method and Mean absolute error(MAE) are found. Results prove that CF-based collaborative filtering achieves smaller MAE when compared with existing methods. The figure below depicts the graphical representation of the Mean absolute error(MAE) values yielded by the existing method and new method.

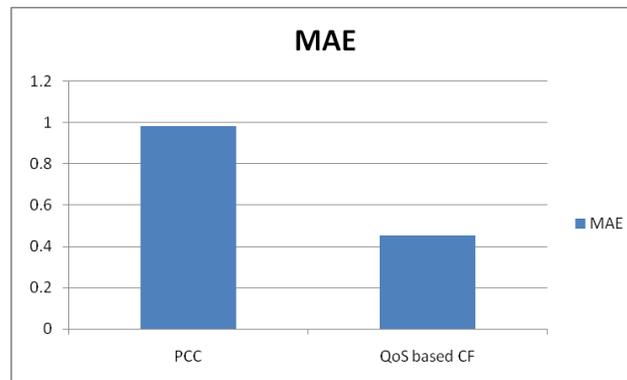


Figure 8 Mean Absolute Error

The graph below is plotted for comparison of computation time of the existing and proposed method employed to predict unknown values

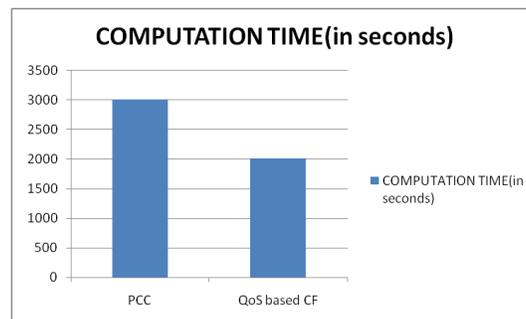


Figure 9 Computation time

Conclusion and future work

Memory based collaborative service recommendation consist of two major steps : similarity measurement and prediction of unknown value. The proposed method produces accurate predictions and gives faster response than the traditional methods.

In future QoS attribute based approach can be applied to other types of collaborative service recommendation systems.

References

- Abinaya, Vinoth Kumar, Swathika, “Ontology Based Public Healthcare System in Internet of Things (IoT),” *Procedia Computer Science.*, vol. 50, pp. 99– 102, Oct. 2015
- B. Vinoth Kumar, M. Ramaswami, P. Swathika, and P. Abinaya, “IPv6 based Patient Monitoring Architecture for Future Healthcare Application” *International Journal of Computer Science and Information Security* Vol. 14, ISSN 1947-5500, Oct. 2016.
- B. Vinoth Kumar, P. Swathika, and P. Abinaya, “Computational Diagnosis Of Parkinson Disease Using Speech Signal” *International Journal of Current Engineering and Scientific Research (IJCESR)* Vol. 03, ISSN (PRINT): 2393-8374, 2016
- Abinaya, P; Aruna; Dharini, Anusha; Devi, Aruna; Devi, Abarna, “Intrusion Detection in Agriculture Field Based on Internet of Things (IoT),” *Journal of Computational and Theoretical Nanoscience*, Vol. 15, Numbers 6-7, pp. 1996-1998(3), June 2018.
- Xiaokun Wu, Bo Cheng, and Junliang Chen, “Collaborative Filtering Service Recommendation Based on a Novel Similarity Computation Method,” in *IEEE trans*, vol 10, No.3, May/ June 2017
- H. Sun, Z. Zheng, J. Chen, and M. R. Lyu, “Personalized web service recommendation via normal recovery collaborative filtering,” *IEEE Trans. Serv. Comput.*, vol. 6, no. 4, pp. 573–579, Oct.-Dec. 2013.
- Y. Cai, H.-f. Leung, Q. Li, H. Min, J. Tang, and J. Li, “Typicalitybased collaborative filtering recommendation,” *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 3, pp. 766–779, Mar. 2014.
- X. Chen, Z. Zheng, X. Liu, Z. Huang, and H. Sun, “Personalized QoS-aware web service recommendation and visualization,” *IEEE Trans. Serv. Comput.*, vol. 6, no. 1, pp. 35–47, Jan.-Mar. 2013
- H. Liu, J. He, T. Wang, W. Song, and X. Du, “Combining user preferences and user opinions for accurate recommendation,” *Electron. Commerce Res. Appl.*, vol. 12, no. 1, pp. 14–23, 2013.\
- Q. Yu, “QoS-aware service selection via collaborative QoS evaluation,” *World Wide Web*, vol. 17, no. 1, pp. 33–57, 2014
- Z. Zheng, H. Ma, M. R. Lyu, and I. King, “QoS-aware web service recommendation by collaborative filtering,” *IEEE Trans. Serv. Comput.*, vol. 4, no. 2, pp. 140–152, Apr.-Jun. 2011
- C. M. Fong, B. Zhou, S. C. Hui, G. Y. Hong, and The Anh Do, “Web content recommender system based on consumer behavior modeling,” *IEEE Trans. Consum. Electron.*, vol. 57, no. 2, pp. 962–969, May 2011