International Journal of Research in Engineering and Science (IJRES)
ISSN (Online): 2320-9364, ISSN (Print): 2320-9356
www.ijres.org Volume 10 Issue 10 ǁ October 2022 ǁ PP. 226-231

Enhancing similarity-Based Web Service Recommendations Using Collaborative Filtering

K.SELVARANI [1] and Dr.A.ANNADHASON[2]
[1] Research Scholar, Department of Computer Applications,
Alagappa University, Karaikudi, Sivagangai (District), Tamilnadu, India
[2]Head &Assistant Professor& HOD,Department of Computer Science,
Govt., Arts &Science College, Kadaladi, Ramnad (District), Tamilnadu, India

Abstract
In the field of the Internet, RecommenderSystem (RS) acts the most significant part in the Data retrieval techniques in request to manage the online data efficiently. it provides recommendations on the items or services to the end user to ensure accurate decisions easily. Torecognize the suitable web service we use some methodologies (Pearson Correlation-Coefficient (PCC) method cored over Collaborative filtering and Novel Recovery-Collaborative filtering (NRCF) method. These methods use the algorithm for similarity measurement and similarity computation in the zone of the Web service recommender system. In contrast with existing Methods, theproposed system has composite Clustering techniques to improve the Similarity measurement Accuracy prediction. To enhance the web service recommendation system efficiency, we conduct the test on PCC over a new composite clustering technique and display the result based on a comparison of the techniques.

Date of Submission: 24-09-2022 Date of acceptance: 08-10-2022

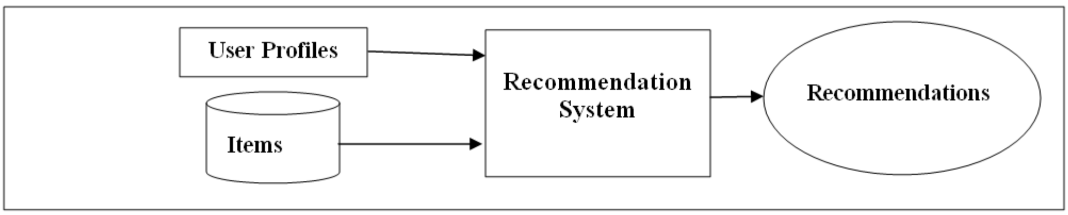
I. Introduction

Web services recommendation systems are the function of automatically finding the needed service and also recommending the web services to the requested user [1]. Nowadays numeric growth of information systems and the rapid increase in the number of users and data have become challenging problems in filtering the needed information. Efficient web service-recommendation is to fulfill the both basic functional and non-functional requirements of the end users. Basic Functional-requirements based upon what a service does and basic non-functional requirements foundation over the Service Quality i.e.) Quality of Service (QoS) like response time, throughput, Round-Trip Time (RTT), etc., the major role in QoS is followed by those same web services that could be ranked and also selected by users [2]. Collaborative filtering is one on one technique that is pre-owned to portend the items that an end user may like on the grounds of ratings provided by other end-users who have the same taste with these items[3][4][5]. When identifying the similarity computation within the end-users and items, most top N-rated recommendation algorithm is pre-owned to refer the Top N-ranked web service to end users.

II. Recommendation System

A web service Recommendation System is one kind of the strongest tools for rising profits and preserving buyers. Recommendation Systems are mostly used in various kinds of fields like E-learning, E-Commerce, Social media, and so on. The recommendation system predicts the interest and preferences of users.

A simple framework of the recommendation system is as follows:



Web service Recommendation system is broadly classified into various types according to the Knowledge principles they use to make recommendations.

❖ **Collaborative Filtering Recommendation:**

One of the most used recommendation systems is collaborative filtering. Collaborative filtering determines the features of items or web services for a particular user to the same in order to the same set of web services given by many other users.

❖ **Content-based Recommendation:**

The content-based recommendation system is recommended items to the user similar to the ones she and he favored in the past. it is adapted to design to recommend text-based items.

❖ **Context-aware Recommendations:**

Context-aware recommendations are the multilateral idea that has been learned beyond various research zone disciplines, computer science, well-known linguistics, philosophy, organizational sciences, and psychology.

❖ **Knowledge-based Recommendations:**

Knowledge-based recommendations apply the knowledge around end users, items, and also which type of products meet users' requirements. An important advantage of Knowledge-based recommendations is the non-existence of cold-start problems.

❖ **Demographic-based Recommendations:**

Demographic-based Recommendations work on the basis of user demographic data. The main aim of the Demographic-based Recommendation approach is to categorize theend user over the base of features and user's demographic information stored in their profiles, for suggesting the item.

❖ **Hybrid Recommendations:**

The hybrid Recommendation approach combines aspects of multiple kinds of filtering approaches by computing the choice of every feature thereby attaining better performance. Hybrid Recommendations merge all the techniques to rise better system optimization.

1. Collaborative Filtering

The process of identifying the same end users and same web services and also recommending what similar users like is called collaborative filtering. The collaborative filtering suggested the web services to theuser, on the base of the past behavior of web service. The end user can hardly invoke all services, meaning that the Quality of Service (QoS) (round-trip time i.e. RTT) ethics of web services that the end-user has not invoked are unknown. Hence, providing accurate Web service recommendation QoS prediction is very important for service users. Table 1 shows a simple example. The numerical ethics in the given table correspond to the response time of the users to invoke the indicated service. Question Mark (i.e.?) means that the end-user has not invoked this service.

Table 1: User-Item Matrix with Round-trip time (RTT) A.

User	Location	Service 1	Service 2	Service 3	Service 4
Sam	Chengdu, CN	30000ms	?	200ms	?
Ram	New York, US	600ms	3300ms	?	3300ms
Bam	New York, US	650ms	2600ms	200ms	?
Tam	New York, US	620ms	2500ms	2000ms	500ms
Lam	Shanghai, CN	1000ms	2500ms	2000ms	5000ms

Similarity Computation Measures:

There are two types of similarity measures, i.e., the functional similarity measure and the nonfunctional similarity measure. Input/output/operation names are usually employed to measure the functional similarity computation within two web services. In our paper, rather than the functional similarity, the focus is on the nonfunctional similarity (QoS similarity).

Given a data set consisting of M web service users and N Web services, the supplication records within users and services can be defined by an M × N matrix, which is called a user–service matrix. An entry in this matrix m, n represents a record of invocation (QoS values, e.g., latency and availability), as demonstrated in Table 1. The similarity is mainly split into two types user-based and Item-based similarity measures.

1) PCC (Pearson correlation-coefficient) [7] was introduced in several recommender systems for similarity computation, then it may be smoothly implemented and can attain high accuracy.

A)User-based Similarity Measures:

In user-based similarity measures, the PCC can be engaged to assess the similarity computation within two users x, y by

$$sim(x,y)=\frac{\sum_{i\in I}\left(r_{x,i}-\overline{r_x}\right)\left(r_{y,i}-\overline{r_y}\right)}{\sqrt{\sum_{i\in I}\left(r_{x,i}-\overline{r_x}\right)^2}\sqrt{\sum_{i\in I}\left(r_{y,i}-\overline{r_y}\right)^2}} \quad \text{.... (1)}$$

Where $I=I_x\cap I_y$ is the bunch of web services enforced by users x, y, and $r_{x,i}$ are the QoS value of item i observed by user x and $\overline{r_x}$ defines the average value of user x over the item in i. From the above equation, rates of the PCC are within the interval of -1 and 1.

B)Item-based Similarity Measures:

The PCC could be engaged to scale the similarity computation within two web services x, y by

$$sim(i,j)=\frac{\sum_{x\in U}\left(r_{x,i}-\overline{r_i}\right)\left(r_{x,j}-\overline{r_j}\right)}{\sqrt{\sum_{x\in U}\left(r_{x,i}-\overline{r_i}\right)^2}\sqrt{\sum_{x\in U}\left(r_{x,j}-\overline{r_j}\right)^2}} \quad \text{..... (2)}$$

Where $U=U_i\cap U_j$ is the group of end users who invoked both web services i,j and $r_{x,i}$ is the QoS evaluate of the web-service i observed by the user and $\overline{r_i}$ defines the average evaluation of the web service i observed through users in Similarly, $r_{x,j}$ is the Qos estimate of the web-service j observed b user and $\overline{r_j}$ defines the moderated rate of the web-service j observed through users in rates of the PCC are within the period of -1 and 1.

2) NRCF

NRCF has been introduced for computing similarity within end users or items (web-services) followed by the user-item value shown in Table which is usually called user-item matrix and prediction of unknown QoS value followed by the similarity computation of users and items.

A)User-based Similarity Measures:
In terms of user-based similarity measures, the PCC can be employed to measure the similarity between two users x and y by

$$Sim(x,y)=1-\frac{\sqrt{\sum_{i\in I}\left(\frac{r_{x,i}-r_{x\min}}{r_{x\max}-r_{x\min}}-\frac{r_{y,i}-r_{y\min}}{r_{y\max}-r_{y\min}}\right)^2}}{\sqrt{|I|}} \quad \text{....(3)}$$

Where $I=I_x\cap I_y$ is the group of web services enforced through users x and y, $|I|$ is also the number of I value, $r_{x,i}$ are the estimates of item i with user x in the original user-item matrix P, $r_{x\min}$ and $r_{x\max}$ defines the lowest and highest values from user x in Respectively, and $r_{y\min}$ $r_{y\max}$ defines the lowest and the highest values from user y in P, respectively. In (3), $Sim(x,y)\in[0,1]$, where $Sim(x,y)=0$ represents that two end users are dissimilar and $Sim(x,y)=1$ indicates that two end users are the same. While the rate of the PCC is in the interval of -1 and 1.

B) Item-based Similarity Measures:

To calculate the correlation between two web services, similarly, based on normalizing the items QoS values, propose a system that guides the standard user-item matrix P into the column-normal user-item matrix $P_{ni,}$ where each column is in the range of $[0,1]$ the method of NRCF is to measure the similarity computation betweenthe two items i,j by

$$Sim(i,j) = 1 - \frac{\sqrt{\sum_{x \in U}\left(\frac{r_{x,i} - r_{i\min}}{r_{i\max} - r_{i\min}} - \frac{r_{x,j} - r_{j\min}}{r_{j\max} - r_{j\min}}\right)^2}}{\sqrt{|U|}} \quad ....(4)$$

Where $U = U_i \cap U_j$ is the bunchof end users who invoked both items i,j and $|U|$ is the numeric of U, $r_{x,i}$ is the scale of item i from user u in the way of original matrix P, $r_{i\min}$ $r_{i\max}$ $r_{j\min}$ and $r_{j\max}$ define the lowest value of item i, the highest rate as to of item i, the lowest rate of the item j, and the foremost value of item j in the initialstandard matrix respectively.

III. Clustering Algorithm

The clustering algorithm chooses a set of similar items or sets of similar users for target users. The Process of selecting a similar item for similar users is crucial for the accuracy of prediction because the prediction of the unknown value depends on the corresponding values of similar users. Propose system uses a clustering methodology.

Clustering is of two-type user-based clustering [7] and item-based clustering.

A) user-based clustering and item-based clustering

User-based clustering is the process of finding a similar neighbor and making a cluster of it. Each user has its own user cluster and this cluster is used for finding unknown QoS values for that respective user. Our proposed system uses the composite clustering algorithm [7] used for user-based clustering. User-based clustering algorithms use user-based similarity values and also Item-based clustering algorithm is used for making a cluster of the same web services. This kind of algorithm is mainly used in NRCF similarity measurement.

Algorithm for the user-based clustering algorithm

Input :

x : Target User $\mathrm{T}(x)$: set of another user

I:Similarity thresold K:Number of similar neighbour to be selected

Output : $S(x)$: Cluster of items for User x.

1.int $\mathrm{N_{sim}} = 0$

2.for all $\mathrm{x_i}$ such that $\mathrm{x_i} \in \mathrm{T}(x)$ do

3.$\mathrm{N_{sim}}{+}{+}$

4.end for

5.If $\mathrm{N_{sim}} \geq K$ then

6. $S(x) \leftarrow Top - K$ user

7. else if 0<$\mathrm{N_{sim}} < \mathrm{K}$ then

8. $S(x) \leftarrow Top\mathrm{N_{sim}}$ items

9. else if $\mathrm{N_{sim}} = 0$ then

10. $S(x) \leftarrow 0$

11. end if

B). Composite clustering

Identifying the QoS values of a user who all has the very lowest similarity to the target user is the useless or very harmful user. So, the Traditional Top-K algorithm is not suited for this problem. Item-based clustering [8] of NRCF similarity computation measurement is not suitable for this kind of problem. So, intending to address this problem, our system proposes a composite clustering algorithm. The composite clustering algorithm is the system proposed technique, which is required for both PCC item similarity computation along with PCC user similarity computation measurement values and gives the output $S(x)$, which means that the group of items for user x.K is the top selected items from the whole group of items, the selection is made by choosing of maximum similarity computation between items.

Below composite clustering algorithm having three threshold values is used. The first value is I, which is used as item similarity computation also which is the lower limit of the qualified similar item. The second value is used as the user Similarity computation measurement threshold, which is the lower limitof qualified users. Finally, the third threshold is limiting the number of users who accessed web service i.

Algorithm for Composite clustering algorithm

Input:

x: Target user W(u):set of web services used by target user u.

$\mathrm{W}(\mathrm{j})$: set of other users web services

I : PCC item similarity threshold U:PCC user similarity threshold

UI:Number of users those used web service i threshold

K:Number of similar items selected

Output:$S(x)$: PCC cluster of items for user x.

1.int $\mathrm{N}_{\mathrm{sim}}=0$

2.for all i such that $\mathrm{i} \in \mathrm{W}(i)$ do

3. count users those access service i into C

4.for all j such that $\mathrm{j} \in \mathrm{W}(j)$ and $\mathrm{sim}(i, j) \geq \mathrm{I}$ and $\mathrm{C} \geq \mathrm{UI}$ do

5. $\mathrm{N}_{\mathrm{sim}++}$

6. end for

7.end for

8.if $\mathrm{N}_{\mathrm{sim}} \geq K$ then

9. for all y such that $\mathrm{y} \in \mathrm{U(y)}$ and $\mathrm{sim(x,y)} \geq \mathrm{U}$ do

10. $S(x) \leftarrow$ web services of y

11. end for

12.else if $0<\mathrm{N}_{\mathrm{sim}} < K$ then

13. $S(x) \leftarrow \mathrm{TopN}_{\mathrm{sim}}$items

14.else if $\mathrm{N}_{\mathrm{sim}} = 0$ then

15. $S(x) \leftarrow 0$

16. end if

IV. Experiment

This section describes the experimental evaluation of our prospective composite clustering techniques pre-owned within the system.

A) Web-service Recommendation QoS-Dataset:

Many kinds of standard datasets are available for experimental uses. They are Planet Lab, and WS-Dream. The Proposed system uses the WS-Dream dataset of testing.

B) Evaluation Metric:

To appraise the QoS value forecastaccuracy, we use the well-known mean absolute error (MAE) metric. The MAE is the standard absolute derivation of predictions to the ground truth values.

The MAE is defined as:

$$\text{MAE}=\frac{\sum_{x,i}\left|r_{x,i}-P\left(r_{x,i}\right)\right|}{N}$$

C) Performance difference of similarity computation measures:

To show the strength of our PCC similarity measure, we correlate it with other similarity computation measures like NRCF similarity Measures. Below table 2 shows the forecast certainty of distinct similarity computation measures. From Table 2, we could see that the best MAE performance for the purpose of PCC is superior to existing NRCF. Compared with existing NRCF, propose PCC approach significantly improves the forecast certainty.

	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
PCC	0.1208	0.0110	0.1036	0.0951	0.0866	0.075	0.695	0.060	0.0524	0.0439
NRCF	0.1357	0.1254	0.1151	0.1048	0.0946	0.0843	0.074	0.0637	0.0535	0.0432

V. Conclusion

In web service recommendation, Pearson correlation-coefficient and Normal recovery collaborative-filtering is used. Our proposed technique approach investigates the characteristics of web service QoS values and proposes a new clustering technique, which is used to find similar users and yields better QoS results. Proposed composite clustering techniques are used of forecasting the unknowns'values. The experimental exhibits that our method significantly improves the QoS value prediction accuracy of the PCC approach.

References

[1]. L.J.Zhang, J. Zhang, "Services Computing: Core Enabling Technology of the Modern Services Industry", Tsinghua University Press,2007.

[2]. Klein, F. Ishikawa, and S. Honiden, "Towards network-aware service composition in the cloud," in Proc. of the 21st ACM World Wide Web Conference (WWW), 2012, pp.959–968.

[3]. Z.Zheng, H.Ma, M.R. Lyu, "Wsrec: A collaborative filtering based web service recommender system," in Proc. of the IEEE International-Conferenceon Web Services (ICWS), 2009, pp.437–444.

[4]. W. Zhang, H. Sun, X. Liu, and X. Guo, "Temporal QoS- aware web service-Recommendation via Non-Negative Tensor-factorization," in Proc. of the 23rdInternational World-Wide-Web Conference (WWW-2014), pp. 585– 596.

[5]. X. Chen, Z. Zheng, Q. Yu, and M. R. Lyu, "Web service-recommendation via Exploiting location and QoS information," IEEE Trans. Parallel Distribution System., vol. 25, no. 7, pp. 1913–1924,2014.

[6]. R.M. Sreenath and M.P. Singh, "Agent-Based Service Selection," J. Web Semantics, vol. 1, no. 3, pp. 261-279, 2003.

[7]. Jian Wu and Liang Chen," Predicting QualityofService for selection-by Neighborhood-based Collaborative-Filtering," IEEE-Transactions on systems, man and Cybernetics-Systems, Vol. 43, No. 2, March2013.

[8]. Sabanaz S. Peerzada, "Web Service-Recommendation using Collaborative-Filtering," International Research Journal of Engineering&Technology-IRJET, Volume: 04 Issue: 06 June2017.

[9]. Sabanaz S. Peerzade, Vanita D. Jadhav, "A Review on Web Service-Recommendation System Using Collaborative Filtering," International-Journal of Computer-Science -IIJCS, Volume 3, Issue 3, March 2015.

[10]. Yali LI, Shanliang PAN, and Xi HUANG," A Kind of Web Service-Recommendation Method Based on Hybrid Collaborative-Filtering," Journal of Computational Information Systems 8: 20 (2012).

[11]. Xia Jingming, Liu Conghui. (2020)", a collaborative Filtering Algorithm Based on user and commodity attribute mining [J]", Modern Electronics Technique, 2020, 43(23):120-123.

[12]. Kousar Nikhath A, Subramanian K (2018), "Conceptual relevance-based document clustering using concept utility-scale", Asian J Sci Res 11(1):22-31.

[13]. Lin Ruyi, He Feng, Zhao Xiolong(2017), "Collaborative Filtering Algorithm based on K-Means Clustering[J]", Computer Fujian,2017,33(12):6-8.

[14]. Li Yanjuan, Niu Mengting, Li Linghui(2019), "A collaborative Filtering Recommendation algorithm based on a bee colony K-means Clustering Model[J]", Computer Engineering and Science 2019,41(06):1101-1109.

[15]. Chigozirim Ajaegbu (2021), "An Optimized item-based collaborative filtering algorithm [J]", Journal of Ambient Intelligence& Humanized Computing, 2021.