Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 4, Issue, 10, pp.207-212, October, 2012

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

AN EFFICIENT APPROACH FOR IMPROVING THE PERFORMANCE OF QoS IN WEB SERVICES

Palanikkumar, D., Nathiya, G. and Lakshmi, M.

Department Of Computer Science and Engineering, Anna University Of Technology, Coimbatore

ARTICLE INFO

Article History:
Received 19th July, 2012
Received in revised form
24thAugust, 2012
Accepted 30th September, 2012
Published online 30th October, 2012

Key words:
Congestion, Routing, Look-Ahead,
Virtual Ring, Proximity.

ABSTRACT

The effective selection and recommendation techniques are needed for increasing presence and adaption of web services in the real world. The collaborative filtering method collects the similar data from other web services and it predicts the current user value. The main assumption of CF is that the users and items have similar behaviors, they will rate or act on other items similarly. Collaborative filtering approach have many challenges, CF algorithms are required highly sparse data. A user collaborative mechanism is used to collect past web service QoS information from different service users. Then based on the collected QoS data a collaborative filtering approach is designed to predict web service QoS values.

INTRODUCTION

The optimal Web services are used for a large number of available Web services. By using the polynomial time method we can evaluate the missing data. Web services have different QoS properties, some properties are user independent and have identical values for different users and some properties are user dependent. Some QoS properties are user dependent and have different values for different users. It is a challenging task to obtain the values of the user dependent QoS properties, because the client side evaluation required the performance of user dependent. In those previous methods Optimal Web service selection and recommendation are difficult to achieve. To avoid these critical difficulties, a collaborative filtering based approach can be used for making personalized QoS value prediction for the service users. The main goal of collaborative filtering algorithm is to provide suggestions on new items or to predict the usability of some particular items based on the user's interests. The most used collaborative filtering methods are user-based approaches and Item-based approaches. These approaches are most useful in modeling characteristics of users and items, so collaborative filtering techniques can be useful in commercial systems.

The user-based approach and item-based approaches can be combined for predicting the QoS values for the current user by taking the historical Web service QoS data from other relevant users and relevant Web services. The relevant service users are defined as the service users have the similar historical QoS experience with the current user. It can be different from the traditional Web service evaluation approach because collaborative filtering approach without requires the real-world Web service invocations; it can predict the user dependent QoS values. Amazon.com is one of the most used web service recommender but the self hosting web services are less expensive when compare to amazon.com web service recommender. In general collaborative filtering is the process for filtering the particular data or patterns by using different techniques involving the collaboration with multiple agents, data sources, and view points and so on. Collaborative filtering techniques can be used in lots of applications. Some of the applications are sensing and monitoring data - for example in mineral exploration, environmental sensing over large areas or multiple sensors, financial data - they include the financial service institutions that integrate many financial sources, or in electronic commerce and web 2.0 applications are focus on user data, etc. The collaborative filtering is a method to make automatic predictions for the particular likings of a user by collecting the similar information or preferences from many users. The Memory based mechanism can use the user rating data to compute similarity between users or items. This can be used for making recommendations. It is an earlier and mostly used mechanism in the commercial systems. It is the easiest method to implement and is effective. Neighborhood based CF and item-based or user-based top-N recommendations are some types of collaborative filtering mechanism. The neighborhood-based algorithm is used to calculate the similarity between two or more users or items and the results can be showed based on the weighted average of all the ratings.

Similarity computation takes most important place in this approach between items or users. Pearson correlation and vector cosine based similarity are some examples for this approach. The user based top-N recommendation algorithm uses the similarity based vector model to identify the k most similar items. After the k most similar users are found, the set of items are identified based on the aggregation of their

Corresponding author: palanikkumard@gmail.com , smartnathiya1@gmail.com , banu175@gmail.com,

corresponding user-item matrices. The Locality sensitive hashing is the most popular method for finding the similar users which implements the nearest neighbor mechanism in linear time. The advantages of this approach are the explain ability of the results, which is an important aspect of recommendation systems. It is easy to create and use. In this approach we can add the new data easily and incrementally. It need not consider the content of the items being recommended; and the mechanism scales well with co-rated items. This approach has the following disadvantages. It mainly predicts the results based on the human ratings. For example the own website user can rank highest ratings for his website and some competitors put lowest rank for some good companies items or products. So this will cause major problem. In that approach it is mainly related with web related items so if any problem occurs in web services or insufficient data the performance will be decreased. It will cause many problems with large data sets. New user or new item could not be handled by this approach. This is one more problem in this approach. The Models will be developed by using data mining, machine learning algorithms that are mainly concentrates on to find the patterns or data sources based on training data. There are many model based CF algorithms are used to find the similarity computation. Some of the approaches are Bayesian Networks, probabilistic latent semantic analysis, latent semantic models such as singular value decomposition, clustering models, multiple multiplicative factor, markov decision process based models etc..

In general most of the models are creating a classification or clustering technique is based on the test set that are used to identify the user. Depends on the types of principal component analysis the number of parameters will be reduced. There are several advantages with this approach. The sparsity problem will be handled well than when compared to memory based algorithm. It improves the prediction performance and also it provides the initial guide for recommendations. The disadvantages of this approach are it is very expensive to build the model. The user must require the tradeoff between the predicted performance and the scalability of that application. By using the reduction models sometimes we lose some useful models. This approach overcomes the limitation of traditional collaborative filtering approaches. It improves the prediction performance. It solves the data sparsity problem and the loss of information problems. Even though it increases the complexity and these are expensive to implement.

Related Work

P.A. Bonatti et al introduced an on optimal service selection is mainly focused to avoid the service matchmaking problems in web services. The automated service selection is based on the quality of service, cost and trust. The composite services are used to model the workflows in this each and every activity potentially based on different service. The optimal service selection is based on a set of service requests has been given then the services offered by other web services, the matchmaking process result, and a numeric measure [1]. J.S. Breese et al provide an empirical analysis of predictive algorithms for collaborative filtering. Several algorithms have been designed for this task, and the techniques are based on correlation coefficient, statistical Bayesian methods and the Vector based similarity calculations [2]. Finally the accuracy is calculated by comparing all those methods. Correlation and Bayesian methods are mainly based on the nature of dataset, application type and the vote availability to make predictions. The hybrid recommender systems survey and experiments is recommended by R. Burkey to provide suggestions effectively [3]. The recommender systems are used to suggest how to purchase each and every item. Recommender systems are commonly used in electronic commerce and information access to provide suggestions effectively. Recommendation is performed by using the following techniques. They are content based, collaborative, knowledge based and some other techniques. The above methods have been combined to improve the performance of the recommender systems.

Collaborative filtering with privacy via factor analysis is used in e-commerce and direct recommendation applications that is proposed by J.Canny. Its main aim is to provide the privacy collaborative filtering scheme to the SIGIR community and it provides the simplified description for each and every protocol. It is mainly focusing on to predict its accuracy, speed and storage requirements [4]. Now a day most systems use centralized databases and these provide privacy risk problems. Collaborative filtering methods reduce these problems'. Cardellini et al introduces the flow based service selection for web service composition supporting multiple QoS classes. A broker that provides a set of services with multiple QoS classes to several users each service have to generate a flow of request over time[5]. A service selection scheme is used to optimize the end-to-end aggregated QoS of all incoming request flow which means a simple linear programming problem that mainly focuses on the number of users involved and the amount of requests has been established. An investigation on personalized collaborative filtering for web service selection that mainly focuses on service discovery problem and not on service selection introduced by K. Karta. The recommender systems are the best way to select the web services [14]. The content based, knowledge based and memory based are some examples for recommender systems. Web services are increased day by day it needs effective software tools to collect the effective information within the scheduled time. In this system we mainly focused on centralized databases only. In the future we have extended this work for distributed system also.

Sarwar et al introduces a new collaborative filtering algorithm based on item based classification. This item based collaborative filtering is based on the classification of items is used to pre produce the ratings. This item based classification is used to classify the items to predict the ratings of those items, and then based on those ratings an item based collaborative filtering is designed to produce the recommendations [19]. In collaborative filtering the sparsity problem is the main problem that is the major reason for poor quality products. The collaborative filtering based on item based classification can avoid the sparsity problem of the user-item dataset. The item based classification can provide better recommendation when compared to traditional collaborative filtering algorithms'. Rosario et al created a probabilistic QoS and soft contracts for transaction based web services orchestrations. In web services service level agreements have an important role and these agreements are used to define the obligations and the rights have been established between the

web service and its client. The soft contract orchestrations are implemented by using the torque tool. In this method we have to monitor continuously any changes has been occur in web services [18]. The variations are based on the response time, availability, maximum allowed query rate and the security. Z. Zheng et al construct a collaborative filtering based web service recommender system. A user contribution mechanism is used in web service recommender systems. It collects the QoS information from other similar users and it provides hybrid collaborative filtering algorithm for QoS value prediction for the web service. The WSRec provides better prediction accuracy when compare to other approaches. In WSRec a systematic QoS information mechanism is used to design the recommender system and it provides the better results. A novel hybrid recommendation method can be used to avoid the data sparsity problem [20]. It has taken both the user based and item based approach that is the main advantage for this collaborative filtering.

Problem Definition

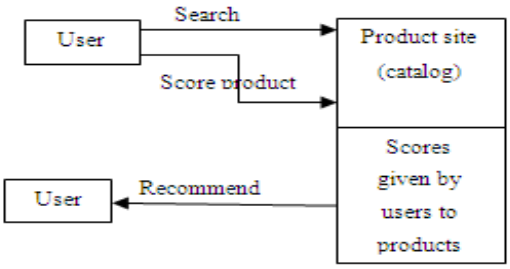

Fig 1.1 Product selection

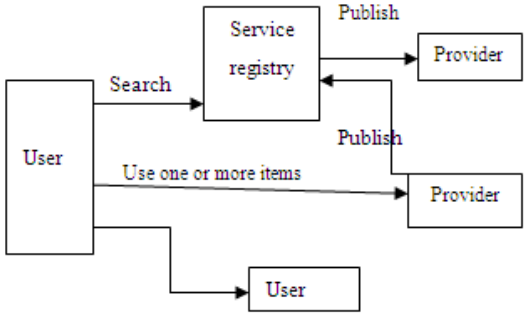

Fig1.2 Service selection

Web services are loosely coupled systems that are mainly used to provide interoperable machine to machine interaction over the internet. By using the collaborative filtering method we can use both the user based approach and the item based approach. The collaborative filtering method collects the similar data from other web services and it predicts the current user value. Hybrid collaborative filtering method can be used to predict the QoS performance of Web services for active service users without requiring the service users because they conduct Web service evaluation and to find out a list of service candidates themselves. Collaborative filtering methods are valuable in e-commerce and for direct recommendation applications. Today most of the companies have been using centralized databases but centralized databases have privacy risk problems. The collaborative filtering method based on factor analysis can be produce best results in speed and storage when compare to other methods.

Implementation Methodology

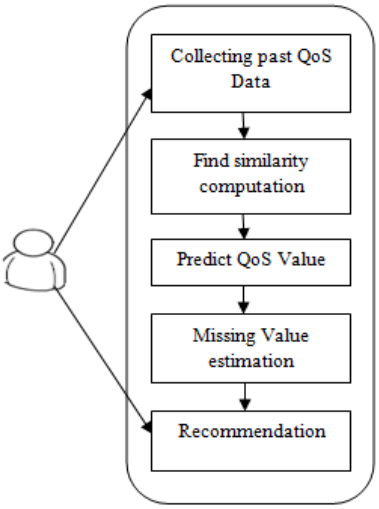

Fig. 1.3. Procedure for QoS Value Prediction

We obtain the list of available web services from internet by searching information from well known web service companies like web service engines and some free web services are hosted in the web. It is difficult to monitor all the web services at the same time, so we have to select some web services randomly they are located in different countries. The selected web services can be replaced for the following reasons. It may require some authentication, sometimes due to some internal problem the system will be shut down and it takes long time duration for processing. By using the collaborative filtering method we can use both the user based approach and the item based approach. The collaborative filtering method collects the similar data from other web services and it predicts the current user value. In collaborative filtering the item based approach is used to pre- produce the user data ratings through the item classification. The collaborative filtering method is used to avoid the data sparsity problem. It is difficult to obtain sufficient information from all the web services. In collaborative filtering we introduce the technique user collaborative mechanism that means contributing the individually observed QoS web service information, so that the users can obtain the accurate recommendation web service.

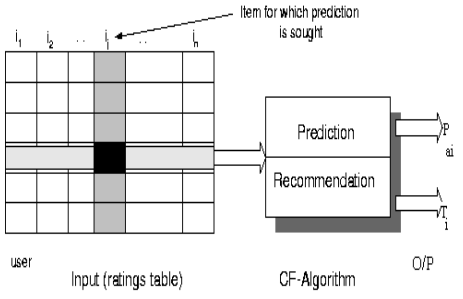

Fig1. 4. Collaborative filtering process

In the above diagram the i_1, i2… i_n are the numbers of items in the web service engine. From that we will extract the relevant items based on the active user's request. Based on the users request the similar items will be predicted. From these there is large number of similar items will be possible. The top 100 similar items is listed from those items. Then the recommendation is performed based on those most similar items. The collaborative filtering algorithm is used to predict the similar items. The similarity computation is performed either weighted average method or mean absolute error method. The mean absolute error is the most commonly used method to find the similarity computation. It will be the average deviation among all those items.

P_{ai} = Prediction item I for the active user

$T_i = T_{i1}$, Ti2, .T_{in} the list of items for the active user

Collaborative filtering algorithm can find the items by using the two methods.

Prediction: It is a numerical value, can be expressed the predicted likeliness of item

Recommendation: It is a list of N items based on the user's likings on those items. The recommended items are not already purchased by the active user.

RESULTS AND DISCUSSION

The collaborative filtering method is based on similar to similar method preferences. The system will collect all the relevant information and then it compares that information with the current user. The performance of the collaborative filtering can be predicted either by using statistical approach or decision support accuracy metrics. First we have to select some web services randomly that they are located in different countries because it is difficult to monitor all the web services at the same time. The final result will be based on how much data is provided by each and every service and how fast that service response for user's request. In statistical approach we use the term Mean Absolute Error (MAE).It is a widely used metrics between the ratings and predictions. It is the measure of deviation among the recommended items. For each and every recommendation we have the prediction pair<A_i,B_i>. It is used to calculate the absolute error between them.

Similarity computation by using MAE method

$$MAE = (\Sigma Ni = 1|Ai - Bi|)/N \qquad (1)$$

The lower the MAE value the recommendation engine predicts the most accurate user item ratings. The decision support accuracy metrics are used to evaluate how effective a service engine is helped the user to select the high quality items from the set of all recommended items. Reversal rate, weighted errors and ROC are some of the decision support accuracy metrics. MAE is the most commonly and easiest method to predict the absolute error between the items. These are the some of the tools used in collaborative filtering algorithm. Memory based tools (CR,VSIM,K-NN,PD, matching), Model based tools (Rules,BC,BN,Boosting), Probabilistic based tools (PD,BN,BC,PLSA,LDA), Distance based (Matching,VSIM,K-NN,CR,PageRank) and ranking based are some of the tools used in collaborative filtering algorithm. First we have to find the list of web pages. Search each artist based on their name and list the top 100 URL's. From these extract the most popularly used websites from all the users. By using the above tools we can find the most used web page. There are lots of websites available in the web service engine. First we have to select the similar website URL from that address. Then based on the URL the service engine will extract the relevant information from that websites. The above tools are mostly used to extract the data from service engine.

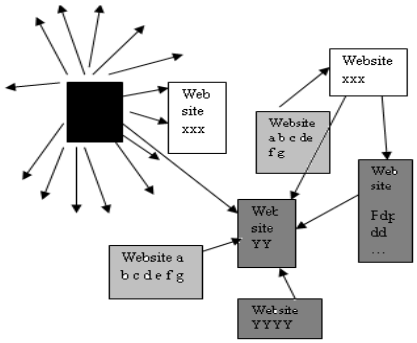

Fig 1.5. An Example for Google's page rank

It is an example for Google's page rank. The Google rank the pages based on how many users visit that particular site. The In links are representing the good site recommendations. Inlinks from a good site are better when compared to inlinks from a bad site. But inlinks from one site with many other outlinks are not good. The good and bad links are relative with one another. It is mainly based on the human ratings for that particular website. It shows the cached or most recently used websites for the active user. For example consider a page hopper. The page hopper that always either follows a random link or jumps to some other pages. The Page rank ranks the pages each and every time by counting how much amount of time the user spent his time in that particular website. If many page hoppers available the page hopper ranks based on the expected crowd size. Some examples of collaborative filtering applications are best sellers list, Top 10 music lists, The recent return books in library, The printers work room, most used weblogs etc..,. From that Inlinks are represent the good item recommendations. The Inlinks are represents the good website. If one website has lots of Inlinks from other websites, the other websites recommend that website is a good one and it contains the more information and it provides the better response time for the active user, and it provides the sufficient information for that particular user.

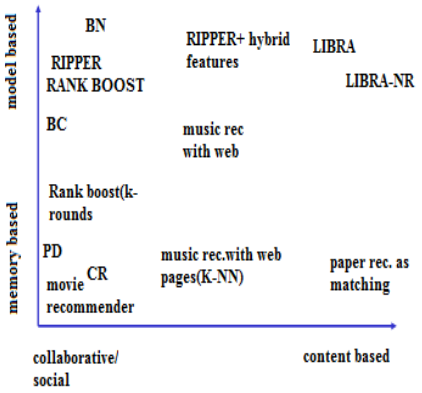

Fig 1.6. Different types of tools used in collaborative filtering approach

The LIBRA book recommender is one type of content based approach. It makes a database to predict the textual description and the Meta information about the particular book from several websites. For example it is based on the Author's name, Book's title, Synopsys, published reviews about the book, user's comments, related authors, related titles etc. These are the different types of approaches used in the collaborative filtering approach.

Prediction computation

The Memory based algorithm uses the following methods to find the similarity computation.

K- Nearest neighbor

$$W(u,i) = \begin{cases} 1, if\ i \in\ neighbors(u) \\ 0 \qquad\qquad else \end{cases} \qquad (2)$$

Pearson correlation coefficient

$$w(u,i) = \frac{\sum_i (v_{a,i} - \overline{\mathrm{V}}\,\mathrm{a})(v_{i,j} - \overline{\mathrm{Vi}})}{\sqrt{\sum_i (v_{a,i} - \overline{\mathrm{V}}\,\mathrm{a})^2 (v_{i,j} - \overline{\mathrm{Vi}}))^2}} \qquad (3)$$

Weighted sum

It is the one more method to compute the similarity items. This method is used to calculate how many users will rate the similar items. We can denote the prediction by using the following formula

$$W(u,i) = \frac{\Sigma_{all\ similar\ rated\ items,N(s,N*\mathcal{R}_{u,\mathbb{N}})}}{\Sigma_{all\ similar\ rated\ items,N(|s,N|)}} \qquad (4)$$

Regression

It is the one more method similar to weighted sum method but in this method it calculates the approximate ratings only. In weighted sum method we directly calculate the similar items instead of using the approximate ratings. The approximate ratings are calculated based on the regression model.

Prediction for active users

After the similarity computation is performed then predicts the Qos values for the active users by applying the user-item matrix for all the items. The prediction is performed based on the response time, how much faster the data is retrieved from the web service engine, how much data will be provided for that web service engine, bandwidth for that items, storage capacity. We predict the QoS performance by calculate the user mean(u) and the item mean(i), where user mean is the average value for different user items and the i-mean is the average value the different items that are observed by the user. It can be denoted by using the following formula

$$\mathrm{P}(M_a, i) = w_u \times \overline{Ma} + w_i \times \overline{Mi} \qquad (5)$$

Evaluation

The evaluation is performed based on the following steps. First the active users can split into two types. First one is test set and the second one is the train set. For each and every user 'a' in the test set we have to split the a's observed votes and then predict P. Measure the average deviation between the observed votes and the predicted votes. Then based on the predicted votes we have to form a Ranked list for each and every user. Then find the average between all the users.

Qos Value Performance Prediction

The missing value prediction can improve the performance of QoS and also it improves the performance of an active user.

Similar Neighbor Selection

We must find the similar items before predict the missing value estimation. Similar neighbor selection is an important process to find the accurate missing values in the items, because the different active users will decrease the accuracy of those particular items. The Top-K recommendation algorithm uses the similar neighbor selection algorithm to find recent used web pages in that items.

Missing value prediction

It is the one more method to find the missing values from the user item matrix.

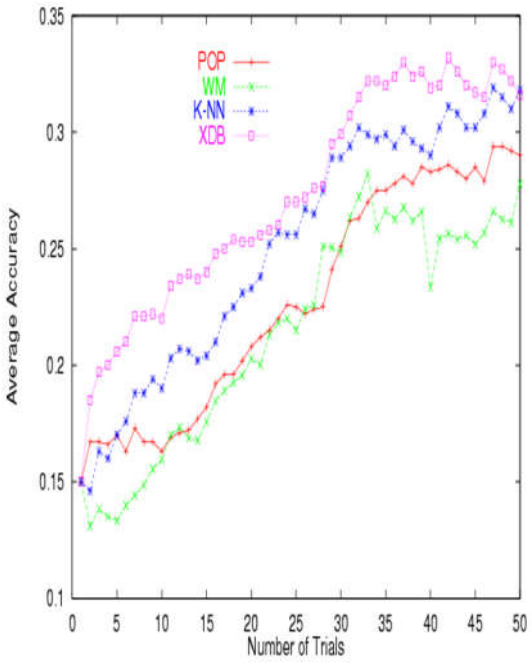

Fig 1.7 Baseline Results for various CF methods

These are the some different types of collaborative filtering algorithm.

POP: It recommends the most popular artist in the world.
WM: Weighted majority. It is the combination of simple experts. For example if a person likes an item A then he will like the item A'.
XDB: If an active user u likes one item i_1, rating i according to R (i.i_1) = $Prob_{u'}$ (likes (u', i) |likes (u, i1))
K-NN: It is a baseline recommendation algorithm. For example a user 'u' and a set of artists 'A_u' which rating (u, a) is already known. Then we pick the similar other users for k. for example u_1, u2, u3…u_k. The other artists will be scored based on the popularity of those artists. Then the final recommendation will be predicted based on the top scoring artists. These are the some collaborative filtering algorithms used to find QoS performance.

In collaborative filtering the inputs are may be true ratings, assumed or implicit ratings, and ratings inferred from web pages. Consider a web page that contains list of items. The most popular websites are extracted from that list of items. Each list of artist will be treated as a user in K-NN. The assumption is made many of the item lists will be related in some similar way. In this paper we mainly focus on the

financial data. Amazon.com is one of the most popular web service recommender that recommends the suggestions to purchase the new items or books. From that website we can purchase whatever we wish but it collects the large amount of information from all the websites its bandwidth and the storage capacity are higher. So it is expensive. For large data collections it is possible but for small data's we can use the self hosting web services. That can reduce the cost, bandwidth and storage. In this paper we used the free web services only. It is an idea if lots of free web services are possible we can reduce the expensive of web services.

Conclusion

Web services are loosely coupled systems that have been used to interact the overall network. The collaborative filtering method collects the similar data from other web services and it predicts the current user value. It is used to avoid the data sparsity problem. It is difficult to obtain sufficient information from all the web services. By using the user collaborative mechanism we can collect the individually observed QoS web service information. This approach will be better when compared to traditional collaborative filtering. It provides quick response time and better storage when compared to other methods. The self hosting web services are less expensive when compare to Amazon.com web service recommenders. In this paper we can use the free web services only. It is an idea if lots of free web services are available we can use that effectively. It will reduce the data sparsity problem and improves the scalability.

REFERENCES

Bonatti, P.A. and P. Festa, "On Optimal Service Selection," *Proc. 14th Int'l Conf . World wide web* (WWW '04), pp. 530-538, 2005.

Breese, J.S., D. Heckerman, and C. Kadie, "Empirical Filtering Analysis of Predictive Algorithms for Collaborative," *Proc. 14th Ann. Conf. Uncertainty in Artificial Intelligence* (UAI'98), pp. 43-52, 1998.

R. Burke, "Hybrid Recommender Systems: Survey and Experiments," *User Modeling and User-Adapted Interaction*, vol. 12, no. 4, pp. 331-370, 2002.

Canny, J. "Collaborative Filtering with Privacy via Factor Analysis," *Proc. 25th Ann. Int'l ACM SIGIR Conf. Research and Development in Information Retrieval* (SIGIR '02), pp.238-245, 2002.

Cardellini, V., E. Casalicchio, V. Grassi, and F.L. Presti, "Flow- Based Service Selection for Web Service Composition Supporting Multiple QoS Classes," *Proc. Fifth Int'l Conf. Web Services* (ICWS '07), pp. 743-750, 2007.

Chun, B., D. Culler, T. Roscoe, A. Bavier, L. Peterson, M. Wawrzoniak, and M. owman, "PlanetLab: An Overlay Testbed for Broad-Coverage Services," ACM SIGCOMM Computer Comm. Rev., vol. 33, no. 3, pp. 3-12, July 2003.

Deora, V., J. Shao, W. Gray, and N. Fiddian, "A Quality of Service Management Framework Based on User Expectations," *Proc. First Int'l Conf. Service-Oriented Computing* (ICSOC'03), pp. 104-114, 2003.

Haddad, J.E., M. Manouvrier, G. Ramirez, and M. Rukoz, "QoS- Driven Selection of Web Services for Transactional Composition," *Proc. Sixth Int'l Conf. Web Services* (ICWS '08), pp. 653-660, 2008.

Herlocker, J.A. Konstan, A. Borchers, and J. Riedl, "An Algorithmic Framework for Performing Collaborative Filtering," Proc. 22nd Int'l ACM SIGIR Conf. Research and Development in Information Retrieval (SIGIR '99), pp. 230-237, 1999.

Hofmann, T. "Collaborative Filtering via Gaussian Probabilistic Latent Semantic Analysis," *Proc. 26th Int'l ACM SIGIR Conf. Research and Development in Information Retrieval* (SIGIR '03), pp. 259-266, 2003.

Hofmann, T. "Latent Semantic Models for Collaborative Filtering," ACM Trans. Information System, vol. 22, no. 1, pp. 89-115, 2004.

Jaeger, M.C., G. Rojec-Goldmann, and G. Muhl, "QoS Aggregation for Web Service Composition Using Workflow Patterns," *Proc. Eighth IEEE Int'l Enterprise Computing Conf.,* pp. 149-159, 2004.

Jin, R., J.Y. Chai, and L. Si, "An Automatic Weighting Scheme for Collaborative Filtering," *Proc. 27th Int'l ACM SIGIR Conf. Research and Development in Information Retrieval* (SIGIR '04), pp. 337-344, 2004.

Karta, K., "An Investigation on Personalized Collaborative Filtering for Web Service Selection," Honours Programme *thesis, Univ. of Western Australia, Brisbane,* 2005.

Linden, G., B. Smith, and J. York, "Amazon.com Recommendations: Item-to-Item Collaborative Filtering," IEEE Internet Computing, vol. 7, no. 1, pp. 76-80, Jan./Feb. 2003.

Ma, H., I. King, and M.R. Lyu, "Effective Missing Data Prediction for Collaborative Filtering," *Proc. 30th Int'l ACM SIGIR Conf. Research and Development in Information Retrieval* (SIGIR '07), pp. 39-46, 2007.

Maximilien, E. and M. Singh, "Conceptual Model of Web Service Reputation," ACM SIGMOD Record, vol. 31, no. 4, pp. 36-41, 2002.

Rosario, S., A. Benvensite, S.Haar, and C. Jard, "Probabilistic QoS and Soft Contracts for Transaction-Based Web Services Orchestrations*," IEEE Trans. Services Computing*, vol.1, no. 4,pp.187-200, Oct./dec.2008

Sarwar, B., G. Karypis, J. Konston, and J.Riedl, "Item-Based Collaborative Filtering Recommendation Algorithms," Proc.10th Int'l Conf. World Wide Web(WWW '01), pp. 439-446, 2007.

Zheng, Z., H.Ma, M.R. Lyu, and I. King,"Wsrec: A collaborative Filtering Based Web Service Recommender System," *Proc. Seventh Int'l Conf. Web Services* (ICWS '09), pp.437-444, 2009

****