Using Fuzzy-Logic in Decision Support System based on Personal Ratings

Hmood Al-Dossari[1], Sultan Alyahya[2]
Information Systems Department
College of Computer and Information Sciences
King Saud University, Riyadh, Saudi Arabia

Abstract—**The decision making process of selecting a service is very complex. Current recommendation systems make a generic recommendation to users regardless of their personal standards. This can result in a misleading recommendation because different users normally have different standards in evaluating services. Some of them might be harsh in their assessment while others are lenient. In this paper, we propose a standard-based approach to assist users in selecting their preferred services. To do so, we develop a judgement model to detect users' standards then utilize them in a service recommendation process. To study the accuracy of our approach, 65536 service invocation results are collected from 3184 service users. The experimental results show that our proposed approach achieves better prediction accuracy than other approaches.**

Keywords—Service recommendation; standard detection; user ratings; predication; fuzzy logic; decision support

I. INTRODUCTION

Different users have different personal standards in evaluating services. Some users might be harsh in their assessment while others are lenient [1]. Therefore, making a generic recommendation to users regardless of their standards underestimates the complexity behind human preferences and thus may result in misleading recommendations.

In the literature, collecting users' ratings after service usage are commonly used for service recommendation [2]–[4]. The most known rating-based technique is the Averaging-All approach [5]. All ratings from previous users of the services are accumulated and the average rating is calculated. The averaging approach is simple; however, it does not consider how personal user standard may affect choosing services. This is because Averaging-All approach neglects the relevance of ratings; irrelevant ratings maybe aggregated, resulting in inaccurate service recommendation.

In this paper, we propose a standard-based approach to assist users select their preferred services. To do so, we develop a judgement model to discover user's standard then utilize the standard to support user's decision in using a given service.

To study the accuracy of our approach, 65536 service invocation results are collected from 3184 service users. The experimental results show that our proposed approach achieves better prediction accuracy than the Averaging-All approach.

The rest of the paper is organized as follows: section II describes the related work of this research. Section III presents a fuzzy-based judgment model used for inferring user standards, Section IV describes the experimental study carried out to test the standard-based service selection approach. The main results are then discussed in section V and finally the paper is concluded in Section VI.

II. RELATED WORK

Several methods in the literature (e.g. [2]–[4]) have adopted rating-based approach. Ratings are personalized; i.e. they depend on a consumer's expectations expressed in his preferences. These preferences capture how important certain aspects of a service are to a consumer [11]. Some solutions (e.g. that consider the subjective nature of ratings rely on the explicit exchange of consumer requests and preferences [12]). They attempt to understand relevant past quality related to service performance by collecting users' feedback in order to assist future service selection.

As we mentioned previously, the Averaging-All is also used where ratings from previous users of the services are accumulated into a single verdict to establish service reputation. Although the simple averaging rating is good enough considering the simplicity of the algorithm design, and the low cost in the system running, it neglects the relevance of ratings; some ratings may be aggregated while they are irrelevant which could result in inaccurate service selection. Therefore, more advanced approaches have been proposed to enhance decision making process.

Collaborative-Filtering methods are widely used in recommender systems [13] [14]. Generally, there are two collaborative-filtering approaches: user-based [15] and item-based [16]. The user-based collaborative-filtering approach defines the similarity between two users based on the services or products they commonly used or bought. The item-based collaborative filtering approach, on the other hand, defines the similarity between the services or products instead of users. Both approaches do not consider how personal user standard may affect choosing services.

Content-based recommendation [17] is a method that filters information based on user's historic ratings on items. The system registers user rating for specific item and links the rating with the attributes of that item. The interest of the user is learned from the attributes of the items he rated. When a new item needs to be evaluated, the system checks if the item has

attributes similar to previous attributes been rated by the same user. The advantage of this method is that the recommendation is based on the individual's historic data rather than taking others' preferences into consideration. However, it overspecializes recommendations because it is based only on the particular user relevance.

Applying context-aware techniques to realize and recommend products or services to the user has gained lots of attentions. Yang et al. [18] [19] develops an ontology-based context model to represent context and utilize the context to assist users in their decision-making. Abbar et al. [20] provide an approach to recommend services using the log files of a user and the current context of the user. To select and recommend services, those approaches either require historical data which are usually not available in practice, or need to predefine the specific reactions on context using rules.

III. A FUZZY APPROACH FOR INFERRING USER STANDARD

In this section, we describe our proposed approach. It consists of two main stages. First, we introduce the judgment model concept. Second, we explain how the judgment model can support users' decision in using a given service.

A. The Judgment Model

Regardless of service domain, the same level of service might be evaluated differently due to the variations of users' standard [1]. The standard is built as a result of users' expectations and past experiences. Some users might be harsh in their assessment while others are lenient. Therefore, users' standards have to be considered in decision-making process.

The proposed approach attempts to discover user's standard based on their past experience which can be extracted from their previous usage of services (i.e. history of ratings). The judgment model consists of two main components: service rating classification and user standard detection. These are described in detail below.

B. Servic Rating Classification

For each service, ratings can be classified into three judgement levels: lenient rating, moderate rating or harsh rating. For illustration purposes, we assume that the rating score range is from 1 to 5 where 1 means harsh and 5 means lenient. For simplicity, in this paper, the ratings are mapped to the judgement level based on the following: 1-2 mapped to "harsh", 3 mapped to "moderate" and 4-5 mapped to "lenient".

C. User Standard Detection

In this section, we introduce our fuzzy logic based on reasoning model for inferring users' standards using past ratings. We use users' historical ratings as an indicator for users' standards. A user standard is formed based on the proportion of the judgement levels.

Generally, for each user we count how many lenient ratings (L), how many moderate ratings (M), and how many harsh ratings (H). We then determine whether his ratings level in each of the three types (lenient, moderate and harsh) are low, moderate or high based on the following formulas:

$$if \; \frac{x}{q} \leq 0.33, then \; the \; level \; x \; is \; low \qquad (1)$$

$$if \; 0.33 < \left(\frac{x}{q}\right) \leq 0.66, then \; the \; level \; x \; is \; average \qquad (2)$$

$$if \; \frac{x}{q} > 0.66, then \; the \; level \; x \; is \; high \qquad (3)$$

Where x is L, M or H and q is the total ratings provided by the user (i.e. q=L + M + H). Fig. 1 shows mapping each judgment level to low, average or high. The judgment levels (e.g. LS, LA and MA) will be explained in details in the next section.

In this study, we use fuzzy inference rules to discover user standard. We define different sets of inference rules where each leads to a certain standard. Generally, there are three main standards: Lenient, moderate or harsh. There are also three sub-levels: low, average or strong. The user standard can be calculated as follow:

User_standard = f(xL, xM, xH)

Where xL is the judgment level for the lenient class, xM is the judgment level for the moderate class, and xH is the judgment level for the harsh class (i.e. the judgment level takes one of the three values: low, average or high).

The output of the previous function is seven different standards as shown in Fig. 1: Lenient Strong (LS), Lenient Average (LA), Moderate Lenient (ML), Moderate Average (MA), Moderate Harsh (MH), Harsh Average (HA) and Harsh Strong (HS). We defined different sets of inference rules to discover user's standards as shown in Fig. 2.
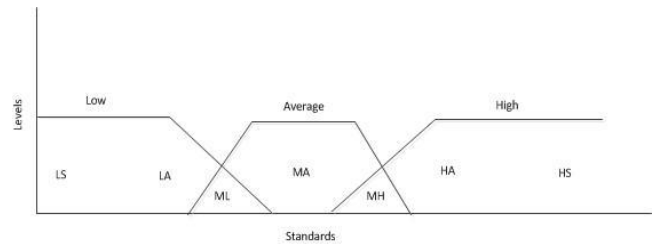

Fig. 1. Mapping Judgment Levels to Low, Average or High.

Let:

xL: the proportion of being lenient

xM: the proportion of being moderate

xH: the proportion of being harsh

Then the fuzzy inference rules are:

Rule 1: If xL *high*, xM *low* and xH *low* then user_standard is ***LS***

Rule 2: If xL *average*, xM *low* and xH *low* then user_standard is ***LA***

Rule 3: If xL *average*, xM *average* and xH *low* then user_standard is ***ML***

Rule 4: If xL *low*, xM *average* and xH *low* then user_standard is ***MA***

Rule 5: If xL *low*, xM *high* and xH *low* then user_standard is ***MA***

Rule 6: If xL *low*, xM *average* and xH *average* then user_standard is ***MH***

Rule 7: If xL *low*, xM *low* and xH *average* then user_standard is ***HA***

Rule 8: If xL *low*, xM *low* and xH *high* then user_standard is ***HS***

Rule 9: If xL *average*, xM *low* and xH *average* then user_standard is ***U***

Rule 10: If xL *low*, xM *low* and xH *low* then user_standard is ***U***

Fig. 2. Fuzzy Inference Rules.

Note that the inference rules in Fig. 2 are incomplete. We have removed impossible situations where L+M+H >1. For example, it is impossible to have xL high, xM high and xH high. On the other hand, there are some situations where it is difficult to clearly classify the user. In such cases the user standard will be classified as unrecognized (U) (Rules 9 and 10).

The basic idea of our approach is to find out the similarity between the consumers based on their standards. When a new consumer would like to use a new service, the approach predicts its suitability based on the ratings of the consumers who have the same level of standard. The proposed approach utilises the discovered users' standard in predicting how likely the user will rate a given service. The details of the approach are shown in Fig. 3.

The algorithm starts by entering current user id (U_id), the potential service (S_id) and the content of the database (DB). The algorithm then retrieves the standard of the current user (α), e.g. whether a user is lenient strong (LS), moderate average (MA) or harsh strong (HS) (Line 1). Then it uses two variables: sum and count, the first for ratings summation (Line2) and the second is to calculate number of ratings (Line3). After that, the algorithm starts to scan the ratings instances in the database (from Line 4 to Line 10). For each instance, it checks the standard of the user (β) and if it is equals to the standard of the target user (Line 6) then its rating will be considered in the predication (Line 7 and 8). After the algorithm finishes from scanning all the existing instances, its predication is calculated by dividing the summation of ratings into the total number of ratings (Line 11) and the result returned to the requester (Line 12).

Note that, in contrast to the average predication method which includes all available ratings about the targeted service, the standard-based algorithm includes only the ratings that were given by a user that has the same standard (Line 6). Such selection will respect the variation between different users in terms of their standards and preferences. An experimental study to expose the efficiency of the standard-based approach is explained and evaluated in the next section.

```
Algorithm  1:  Pseudo-Code  for  Standard-based

Input: U_id, S_id, DB

Output:R

Process:
1.  Read(α);
2.  sum = 0;
3.  count = 0;
4.  For each r on S
5.  Begin
6.  If β = α then
7.  sum = sum + r;
8.          count = count + 1;
9.  End if
10. End for
11. R = sum / count;
12. Return R;
```

Fig. 3.   Standard-based Prediction Algorithm.

IV. EXPERIMENTAL STUDY

In our study, we used MovieLens dataset to test the standard-based approach. It can be downloaded from the GroupLens Research Project website [6]. MoveLens is an experimental platform for studying recommender systems. MovieLens data sets were collected by the GroupLens Research Project at the University of Minnesota. The data was collected through the MovieLens website (movielens.edu). This data set consists of 65536 ratings (1-5) from 3148 users on 438 movies.

The data of all users who used less than five services has been removed due to difficulty of detecting their standards. This has reduced number of users to 2177 users.

A. Experiment 1: Users' Standard Detection

The goal of the first experiment is to detect the standard of each user. For each rating instance, we mapped the rating value to lenient, moderate or harsh as explained in section II. Then, for each user we calculate how many lenient ratings (L), how many moderate ratings (M), and how many harsh ratings (H). Afterwards, the fuzzy inference rules in Fig. 2 are used to detect the user standard. The results are shown in Fig. 4.

As shown in Fig. 4, the majority of the users covered in our experiment have been classified in the Lenient standard classes (LS and LA). Although such result may seem to be surprising, many social studies state the most people are lenient when evaluating specific types of service (e.g. hotels in [7]).

On the other hand, low percentage of users have been grouped in the Harsh standard classes (HA and HS). Specifically, it is clear that number of users who defined as a harsh average is very low (less than 1%). It is hard to justify the reason behind this low level and more investigation might be required. Note that 12.14% of the covered users have not been assigned to any standard classes (i.e. unrecognized).

B. Experiment 2: Standard-based Prediction

In this experiment, we exclude some data out of the dataset, and then use the remained data to predict the excluded. We randomly excluded 30 rating instances from the dataset. Then we compare the prediction of our proposed approach with the Averaging All approach [5]. The prediction of Averaging All approach is calculated as the sum of ratings divided by the number of ratings.
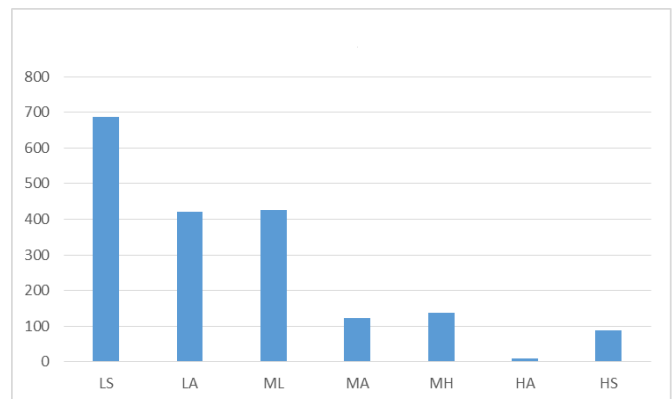

Fig. 4.   User Standard Detection.

The accuracy of each approach is calculated based on how far its prediction from the actual rating. More specifically, the accuracy is calculated based on the following equation:

$$Accuracy = \left(5 - abs\left(\frac{Actual}{predicted}\right)\right) * 20 \quad (4)$$

The result of this experiment is shown in Fig. 5.

Fig. 6 illustrates the result produced by our approach against the result of average prediction method. It shows the comparison of standard-based prediction (our approach) against the Prediction of Average Method. In the Fig. 6, accuracy T represents the accuracy of average method prediction and accuracy S illustrates the prediction of standard-based approach. In the figure, the x-coordinate represents the standard classes as explained in Section 2.1.2 and each has two columns one for the accuracy of our proposed approach and the other for the accuracy of the average method prediction. Y-coordinate, on the other hand, illustrates the accuracy percentage.

From this figure, it is clear that our approach performs much better than the average method in predicting the actual rating for all kind of classes. In harsh standards (HS and HA), our approach outperforms the average method by almost 21%. It is noticed that the accuracy of our approach is similar to the accuracy of the average method in predicting the rating of the lenient users. This can be explained by having large proportion of lenient users (e.g. 31%, 20%, and 20% for LS, LA and ML, respectively). Having such a large proportion makes our proposed approach used almost the same ratings as the average method.

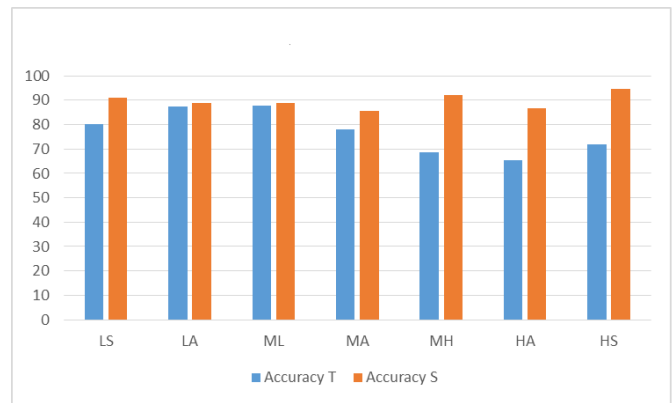| user_id	service_id	Actual_rating	class	Predicted_T	Predicted_S	Accuracy_T	Accuracy_S
2002	424	3	ML	3.73573	3.59193	85.2855	88.1614
586	424	3	ML	3.73573	3.59193	85.2855	88.1614
2058	438	4	ML	3.55098	3.47727	91.0195	89.5455
3863	195	3	ML	3.88929	3.64865	82.2141	87.027
1375	424	4	ML	3.73573	3.59193	94.7145	91.8386
1831	302	1	MH	2.86929	1.6338	62.6142	87.3239
435	48	2	MH	3.06856	2.04651	78.6288	99.0698
173	424	2	MH	3.73573	2.5	65.2855	90
2054	352	3	MA	3.47931	2.83673	90.4138	96.7347
2616	202	4	MA	3.24925	3.26471	84.9851	85.2941
2717	411	2	MA	3.28031	2.6	74.3937	88
3257	302	1	MA	2.86929	2.46341	62.6142	70.7317
2688	319	3	MA	3.26352	2.7	94.7295	94
3785	195	2	MA	3.88929	3.06667	62.2141	78.6667
2858	424	5	LS	3.73573	4.33486	74.7145	86.6972
2396	216	4	LS	3.23174	3.67119	84.6348	93.4237
1196	195	5	LS	3.88929	4.45205	77.7859	89.0411
1210	302	5	LS	2.86929	3.94118	57.3858	78.8235
480	352	4	LS	3.47931	4.16732	89.5862	96.6537
2028	329	4	LS	2.72644	3.37736	74.5288	87.5472
2571	245	3	LS	2.67147	3.32549	93.4293	93.4902
1198	319	4	LS	3.26352	3.96207	85.2705	99.2414
3578	53	5	LS	4.23684	4.76763	84.7368	95.3527
1183	149	4	LA	3.94088	3.78125	98.8176	95.625
2949	173	3	LA	3.96791	3.575	80.6417	88.5

Fig. 5.   Experiment Result.

Fig. 6.   A Comparison between the Result of the Standard-based Approach and the Result of the Averaging-All Approach.

Note that, in the strong classes (HS and LS) the predication of our proposed approach is much better than the average predication method. This is because the standard of those classes are a little bit clear compared to other classes (e.g. LA and ML).

Overall, the accuracy of the standard-based predication outperforms the average predication by almost 13%. While the accuracy of the standard-based was 89.75%, the average predication's accuracy was 77.09%.

V. DISCUSSION

Our approach is based on the assumption that the consumers that have similar standards would have similar feedback (i.e. ratings) on the services. The mapping between ratings and standards was absolute regardless of the range of the ratings. That is, we mapped 1 to "harsh" standard and 5 to "lenient" regardless of its relatives to other ratings. In some cases, for example with excellent services, the range might be limited in a range from 3 to 5. The mapping should respect such case by considering the range in rating-standard mapping. That is, in the range from 3 to 5 for example, it is more reasonable to map 3 to "harsh" rather than to "moderate" standard. On the other hand, if the ratings of a service ranged from 1 to 3, rating 3 in this scenario would be mapped to "lenient" standard. Hence, the rating-standard mapping will be conducted relatively rather than absolute.

The dynamic change of standard over time must be considered. It is possible for a user to switch from being a lenient to be harsh or vice versa. This may happen as a result of having new experiences. Such issue needs to be studied carefully in order to observe and capture any changes in the users' standards.

Additionally, number of required ratings to define user's standard is an essential issue. A study in [8] suggests using the Chernoff Bound theorem [7] to determine the minimum number of ratings required by a given acceptable level of error and a confidence measurement. For example, if we require 80% confidence with 0.2 level of error in determining the user's standards, we need at least 29 ratings from the target user.

A cold start problem [9] [10] is a common problem with any recommender system and the standard-based recommender

is not exceptional. The standard-based approach depends on user's standard to produce a recommendation for the user. With new users, unfortunately, it is hard to suggest any recommendation as their standards are not known. It is quite interesting to think about other types of data rather than ratings to discover users' standards. In eBay (www.ebay.com), for example, the user is required to not only provide their ratings about their experience, but also give their text feedback as a more explanation for their negative or positive ratings. Such text can be utilized to discover their standards in addition to their ratings.

A possible extension for the standard-based approach is to consider the confidence on users' standard detection. A method proposed in [8] can be used to determine the minimum number of ratings needed to be confident about the user standard. Also, it is important to consider how to discover the standard of a new user (i.e. a user has not rated any service yet).

VI. CONCLUSION

In this paper, we have proposed a standard-based recommender system. Because understanding human preferences are complex and may depend on their personal standards, these standards must be considered in the recommender system. To do so, we have developed a fuzzy approach to inferred user standard then utilized the standard to recommend a service to a given user. An empirical study has been conducted using a dataset that consists of 65536 ratings from 3148 users on 438 movies to evaluate the accuracy of our proposed approach against the average prediction method. The result shows that our proposed method has significantly improved the prediction accuracy by almost 13%, compared to the average predication method. The accuracy of our proposed method was 89.75% while the average method accuracy was 77.09%.

ACKNOWLEDGMENT

REFERENCES

[1] P. Lyons, "Performance Management: Concepts, Skills, and Exercises." Academy of Management Learning & Education, vol. 5, no. 2. pp. 246–248, 2006.

[2] H. Liu, L. Guo, L. Chen, X. Liu and Z. Zhu, "Effective Similarity Measures of Collaborative Filtering Recommendations Based on User Ratings Habits," 14th International Conference on Semantics, Knowledge and Grids (SKG), IEEE, 2018.

[3] K. Bauman, B. Liu and A. Tuzhilin, "Aspect Based Recommendations: Recommending Items with the Most Valuable Aspects Based on User Reviews,", in Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, IEEE, pp. 717-725, 2017.

[4] L. Li , Y. Zhou, H. Xiong and C. Hu, X. Wei, "Collaborative filtering based on user attributes and user ratings for restaurant recommendation.", in 2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), IEEE, 2017.

[5] Z. Chen and C. Liang-Tien, "UX-an architecture providing QoS-aware and federated support for UDDI," Proc. …, pp. 1–6, 2003.

[6] "MovieLens." [Online]. Available: http://www.grouplens.org/node/12.

[7] J. Zhang, "Information Revelation and Social Learning," Int. J. Bus. Soc. Sci., 2014.

[8] C. Zhang, Jie, "A personalized approach to address unfair ratings in multiagent reputation systems," in Proceedings of the AAMAS Workshop on Trust in Agent Societies, 2006.

[9] J. Weim J. He, K. Chen, Y. Zhou and Z. Tang, "Collaborative filtering and deep learning based recommendation system for cold start items.", 2016 IEEE 14th Intl Conf on Dependable, Autonomic and Secure Computing, 14th Intl Conf on Pervasive Intelligence and Computing, 2nd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSci Tech), IEEE, 2016.

[10] W. Zhang and J. Wang, "A Collective Bayesian Poisson Factorization Model for Cold-start Local Event Recommendation.", KDD '15 Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, PP. 1455-1464, ACM, 2016.

[11] F. Klan and B. K.-R. Klan, "A personalized approach to experience-aware service ranking and selection," calable Uncertain. Manag., 2008.

[12] M. Şensoy, F. Canan Pembe, H. Zirtiloğlu, P. Yolum, and A. Bener, "Experience-based service provider selection in agent-mediated E-Commerce," Eng. Appl. Artif. Intell., vol. 20, no. 3, pp. 325–335, 2007.

[13] H. Ma, I. King, and M. R. Lyu, "Effective missing data prediction for collaborative filtering," Annu. ACM Conf. Res. Dev. Inf. Retr., pp. 39–46, 2007.

[14] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "GroupLens : An Open Architecture for Collaborative Filtering of Netnews," in Proceedings of the 1994 ACM conference on Computer supported cooperative work, pp. 175–186, 1994.

[15] J. S. Breese, D. Heckerman, and C. Kadie, "Empirical analysis of predictive algorithms for collaborative filtering," Proc. 14th Conf. Uncertain. Artif. Intell., vol. 461, no. 8, pp. 43–52, 1998.

[16] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," Proc. 10th …, vol. 1, pp. 285–295, 2001.

[17] M. J. Pazzani and D. Billsus, "Content-Based Recommendation Systems," Adapt. Web, vol. 4321, pp. 325–341, 2007.

[18] I. Y. L. Chen, S. J. H. Yang, and J. Zhang, "Ubiquitous provision of context aware Web services," in Proceedings - 2006 IEEE International Conference on Services Computing, SCC 2006, pp. 60–68, 2006.

[19] S. J. H. Yang, J. Zhang, and I. Y. L. Chen, "A JESS-enabled context elicitation system for providing context-aware Web services," Expert Syst. Appl., vol. 34, no. 4, pp. 2254–2266, 2008.

[20] S. L. Abbar Sofiane, Mokrane Bouzeghoub and Stephane Lopez, "Context-aware recommender systems: A service-oriented approach.",", in VLDB PersDB workshop., 2009.