Improving Memory-Based Collaborative Filtering by Neighbour Selection based on User Preference Overlap

Alejandro Bellogín[a,b], Pablo Castells[a], Iván Cantador[a]

[a] Universidad Autónoma de Madrid, Departamento de Ingeniería Informática, 28049 Madrid

[b] Centrum Wiskunde & Informatika, Science Park 123, 1098 XG Amsterdam

{alejandro.bellogin, pablo.castells, ivan.cantador}@uam.es[a]; alejandro.bellogin@cwi.nl[b]

ABSTRACT

User-based collaborative filtering approaches suggest interesting items to a user relying on similar-minded people referred to as neighbours. While standard approaches select neighbours based on user similarity, others rely on aspects related to user trustworthiness and reliability. We investigate the extent to which user similarities are essential to obtain high quality item recommendation, and propose to select neighbours according to the overlap of their preferences with those of the target user. We empirically show that a neighbour selection strategy based on preference overlap achieves better performance than similarity- and trust-based selection strategies, in terms of both recommendation accuracy and precision.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval] *information filtering*

General Terms

Algorithms, Measurement, Performance, Experimentation.

Keywords

Collaborative filtering, neighbour selection, user similarity, user preference overlap.

1. INTRODUCTION

Memory-based collaborative filtering represents a type of recommendation approaches that (in its user-based variants) explicitly seeks people – commonly referred to as neighbours – who have preferences – usually expressed in the form of numeric ratings – in common with a target user, and use these preferences to predict item ratings for that user. These approaches are based on the principle that a particular user's ratings are not equally useful to other users as input for providing personalised item suggestions [4]. In this context, main aspects of these approaches are a) how to identify which neighbours form the best basis to generate recommendations, and b) how to properly exploit the preference (rating) information acquired from them. In general, once the target user's neighbours are selected, the more similar a neighbour is to the target user according to their rating profiles, the more her preferences are taken into account as input to provide recommendations. This is usually formulated for a target user u and item i as follows [10]:

$$\tilde{r}(u,i) = \bar{r}(u) + C \sum_{v \in N_k(u,i)} sim(u,v)\big(r(v,i) - \bar{r}(v)\big) \qquad (1)$$

OAIR ´13, May 22–24, 2013, Lisbon, Portugal.

where $\tilde{r}$ denotes a rating prediction (as opposed to observed ratings denoted by r), $\bar{r}(u)$ is the user u's average rating, C is a normalisation factor, and $N_k(u,i)$ is a neighbourhood of size k, which may depend on any or both u and i.

As we may observe, in the above equation the dependency on a user similarity function and a neighbour selection strategy is explicit. Nonetheless, a user's neighbourhood is usually formed with the users who are most similar to the target user (according to their rating profiles), and thus, the neighbourhood and the rating prediction at large only depend on the similarity function used [2]. This has motivated the investigation of better user similarity functions and neighbour weighting schemes. For instance, in [11] Said et al. analyse the effect of neighbour weighting schemes on different types of users, such as users with few ratings (cold start users), users with a bit more than a few ratings (post cold start users), and users with many ratings (power users).

Standard user-based collaborative filtering approaches aim to automatically compute the weights to assign to each pair of users on an item basis [6], [7], [8]. Trust-aware approaches, on the other hand, exploit additional factors based on the concept of trust (trustworthiness, reputation) on a user's contribution to the computation of recommendations [9], [5], [3].

In this paper we focus on the neighbour selection problem. Instead of presenting further variations of user similarity metrics, we fix this component, and explore different neighbour selection strategies to analyse the importance of selecting good neighbours for recommendation. In particular, we propose a number of strategies based on the overlap between rating profiles of the target user and neighbours. Specifically, we address the following research questions: in memory-based collaborative filtering, **RQ1)** is the rating overlap a good surrogate for user similarity? and **RQ2)** can we use it to improve the accuracy and precision of recommendations? For such purpose, we report empirical results on a public dataset where the proposed approach shows noticeable improvements in terms of both accuracy and precision.

2. NEIGHBOUR SELECTION IN USER-BASED COLLABORATIVE FILTERING

To analyse the neighbour selection problem in user-based collaborative filtering, we consider different formulations to define the set of neighbours $N_k(u,i)$ for rating prediction, as in Equation (1). The most common approach consists of a top-N **similarity-based filtering** [2] in which the k users who have the largest similarities with the target user are selected as her neighbours. Other approaches, usually referred to as **trust-based filtering** approaches [9], only consider the most trustworthy users in the recommendation prediction process. In these approaches the neighbour selection can be generically expressed as follows:

$$N_k^{\text{trust}}(u,i) = \{v \in \mathcal{U} \colon \text{trust}(v,i) > \tau\} \qquad (2)$$

where $N_k^{\text{trust}}(u,i)$ denotes the neighbour set of a user u on a specific target item i, for a specific trust metric denoted as $\text{trust}(u,i)$ and a predefined trust threshold τ. In this case the k users with the highest trust values are selected as neighbours.

In this paper we propose an **overlap-based filtering** in which the users who have more preferred items in common with the target user are selected as neighbours. The idea of considering the preference overlap between neighbours as a measure of neighbour appropriateness has been explored in the literature either as a part of the similarity function, or as a post-similarity stage [4], by weighting the participation of the neighbours in the recommendation computation. Differently from these approaches, we investigate the consideration of the above principle as the single criterion for neighbour selection, after which the overlap is no longer taken into account (neither in the user similarity function, nor in any posterior user weighting).

The assumed hypothesis of a preference overlap criterion is that a neighbour with more preferred items in common with the target user is more likely to be reliable. Specific functions to convert degrees of overlap into weights have been proposed in the literature [4][8], such as:

$$\begin{aligned}\text{Herlocker's weighting}(u,v) &= \frac{\min(|\mathcal{I}_u \cap \mathcal{I}_v|, n)}{n} \\ \text{McLaughlin's weighting}(u,v) &= \frac{\max(|\mathcal{I}_u \cap \mathcal{I}_v|, n)}{n}\end{aligned} \qquad (3)$$

These overlap functions were proposed to devaluate a similarity weight when it is based on a small number of co-rated items by two users, that is, when there is a small intersection between the items rated by u and v, denoted respectively as $\mathcal{I}_u$ and $\mathcal{I}_v$. In this context, the parameter n establishes the threshold of what is considered as a "small" (not significant) overlap. In addition to these two weighting functions, we shall consider the intersection size $|\mathcal{I}_u \cap \mathcal{I}_v|$ as an additional, simple alternative.

Overlap functions are commonly used as an element for neighbour weighting in user-based collaborative filtering [4]. Differently from previous work, we propose to circumscribe the use of these weights to (and as the single criterion for) the neighbour selection stage. In this way we filter the neighbours entirely without any similarity function, which provides an advantage in terms of memory and time efficiency, since similarity functions are typically more expensive than the overlap-based strategies presented in Equation (3) [2].

3. EXPERIMENTS

We have tested the proposed overlap-based neighbour selection strategy and compared it to similarity-based and trust-based alternatives. Specifically, we test the overlap method against a recommendation approach that selects neighbours according to their similarity (*similarity filtering*) as in Equation (1), and uses the Pearson's correlation as the user similarity function $sim(u,v)$. We also compare to a *trust filtering* approach in which, as proposed in [9], the parameter τ of Equation (2) is set to the mean across all the trust values for the pairs of users (which resulted in a value of 0.74), and the trust function is:

$$\text{trust}(u,i) = \frac{|\{(v,j) \in \text{CorrectSet}(u) : j = i\}|}{|\{(v,j) \in \text{RecSet}(u) : j = i\}|} \qquad (4)$$

In our experiments the parameter n in Equation (3) was set to 50 unless otherwise stated.

We used the well-known dataset MovieLens 1M, which has one million ratings for 3,900 movies by 6,040 users. We measure recommendation accuracy by the Root Mean Squared Error (RMSE), and recommendation ranking quality by Precision at 10 (P@10). We checked further metrics such as the Mean Absolute Error, the normalised Discounted Cumulative Gain, and the Mean Reciprocal Rank 5, obtaining similar results than those we report in this paper, whereby we omit them here.

3.1 Relation Between Similarity and Overlap

In order to assess the ability of the overlap functions – defined in Equation (3) – to capture the similarity between users, we first compare in Figure 1 the Pearson's similarity values of each pair of users against their overlap values (i.e., the sizes of the intersections of their profiles).

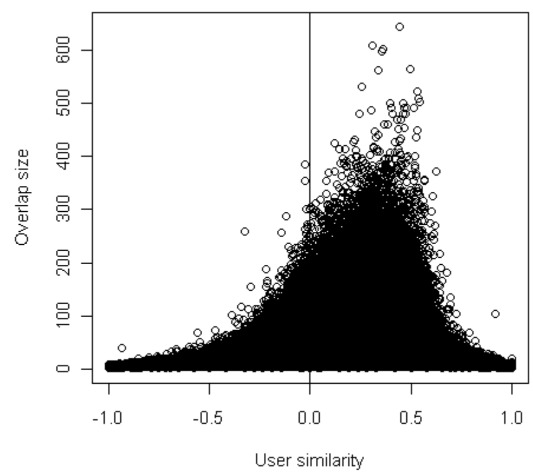

Figure 1. Scatterplot of the MovieLens1M users' similarities (by Pearson's coefficient) against their rating overlap.

The figure shows that the larger the overlap between two users, the more likely such users would agree in their ratings since they tend to have a higher similarity value. Furthermore, the figure also shows that low (negative) similarity values are mostly obtained for users whose rating profiles overlap in fewer items, evidencing that there is some connection between the amount of items a pair of users have rated in common and the similarity between them. Finally, we see that the highest similarity values are obtained by users with a tiny overlap. This is because it is very rare for two users to have almost coincidental preferences in a larger sample of their ratings, whereas with very few ratings, a similarity close to 1 may be observed by chance. This also means that such highest similarities have a very low reliability.

This result provides a positive answer to our first research question **RQ1**, showing that user rating overlap is a reasonable surrogate for user similarity. Our observation moreover suggests that rating overlap could be more robust than common similarity measures to misleading extreme coincidences of ratings on small profile intersections.

This motivates our approach to use the user preference overlap instead of the user similarity function as the neighbour selection strategy in user-based collaborative filtering, as we explain in the next subsection.

3.2 User Preference Overlap for Neighbour Selection

In a second experiment we analyse the recommendation performance of the neighbour selection strategies presented in Section 2. Figure 2 shows the RMSE (left) and P@10 (right) values obtained for a varying number k of neighbours. Specifically, we compare the performances of the similarity and trust filtering strategies against those based on user preference overlap (intersection, Herlocker, and McLaughlin filtering).

The figure shows that the overlap-based filtering approach that uses the McLaughlin's weighting is consistently the best in terms of both RMSE and P@10. Moreover, its performance is better when smaller neighbourhoods are used, that is, it is able to get better results using more *economic* neighbourhood sizes. This is a relevant result because the computational cost of large neighbourhoods is one of the well-known problems of collaborative filtering in terms of computation time and memory [2]. We should note that in our experimental configuration, the coverage (number of users the method is able to recommend at least one item) drops considerably as neighbourhoods get smaller (e.g. down to a 1% at the peak of McLaughlin's accuracy for $n = 100$ with 20 neighbours). This results in an extra improvement of absolute metric values, since uncovered users are the sparsest ones, hence more difficult to predict, which explains some unusually low RMSE values. However this affects Herlocker and McLaughlin filtering alike (actually less so for the latter), whereby the comparative observations remain fair.

The figure also shows that the other overlap-based filtering approaches perform better than trust-based filtering, especially with very small neighbourhoods. This improvement fades down to the baseline when larger neighbourhood sizes are used. Moreover, the overlap-based filtering approaches outperform the baseline (similarity-based filtering) by up to 13% in terms of accuracy, and almost in a factor of 10 in terms of precision (at the expense of coverage, as mentioned above). Neighbour selection by rating overlap thus improves the recommendation performance, providing a positive answer also for our second research question **RQ2**.

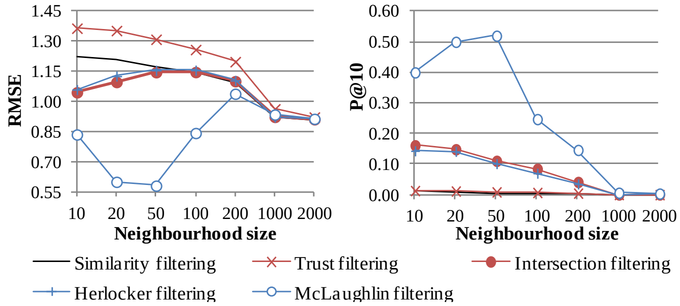

Figure 2. Performance comparison of the evaluated neighbour selection strategies for different neighbourhood sizes.

To further analyse the sensitivity of the approaches to different values of the parameter n from Equation (3) – that is, the threshold that determines when an overlap in the preferences of a pair of users is significant –, we plot in Figure 3 the RMSE and P@10 values for different neighbourhood sizes and values of n (note that $n = 50$ in Figure 2). We can observe that the trend is very similar with different values of that parameter, ranging from 10 to 100 in our experiments.

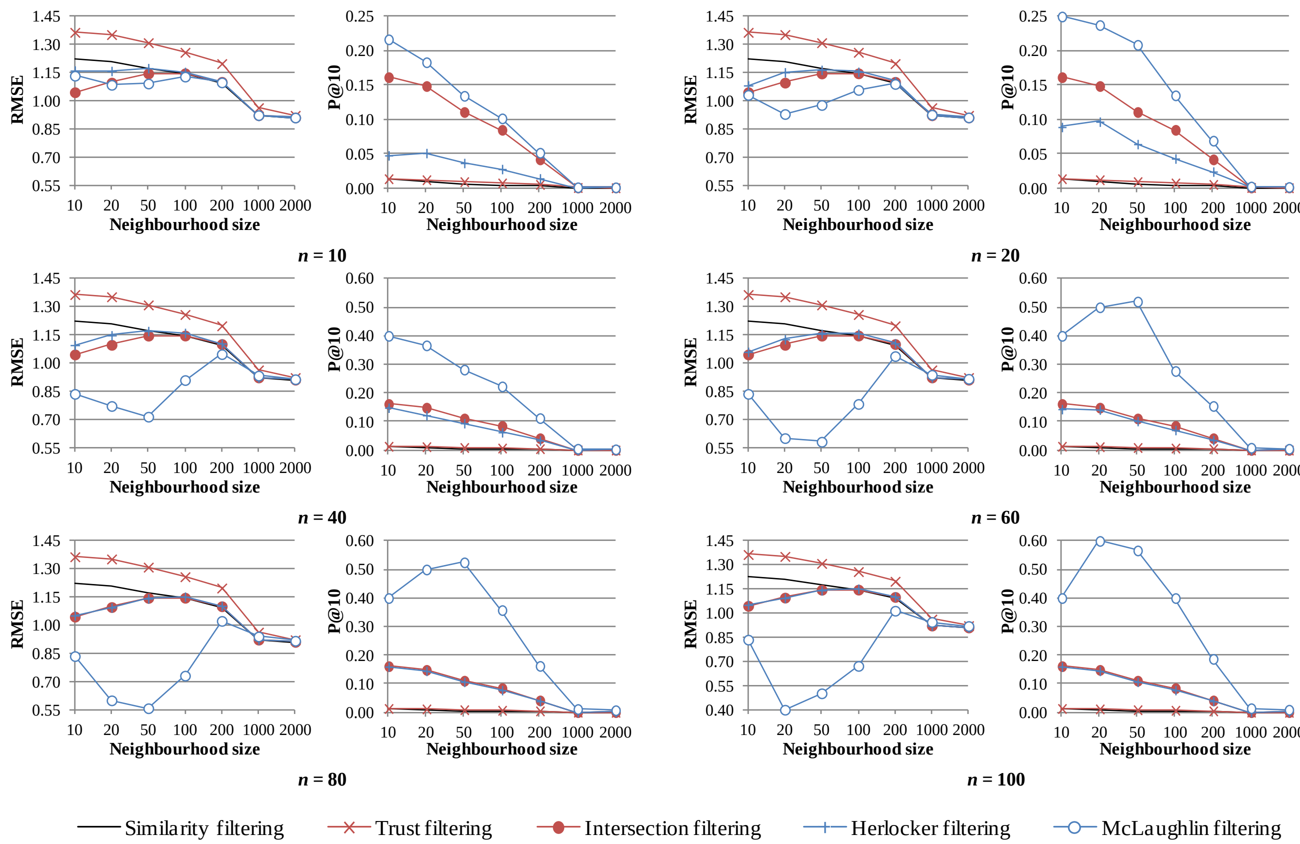

Figure 3. Performance comparison of the overlap-based neighbour selection methods for different neighbourhood sizes, and values of the parameter n related with the significance of the user preference intersections.

We have to note that McLaughlin's filtering performs better when larger values of this parameter are used, although they remain stable for sufficiently large numbers ($n \geq 80$ in our case). On the other hand, Herlocker's filtering degrades to Intersection filtering for large values of n, and thus, its performance decreases. The reason for this behaviour is analysed in Figure 4, where we show the probability mass function of the overlap values shown in Figure 1. We may observe that most of the user pairs have less than 80 items in common, which by definition (see Equation 3) makes the Herlocker's weighting function equivalent to the intersection strategy in most of the cases. Thus, these two strategies are selecting the neighbours only according to the amount of overlap, although in the Herlocker's function, such value is scaled by a factor of n, producing a ranking of the potential neighbours that is equivalent for both strategies.

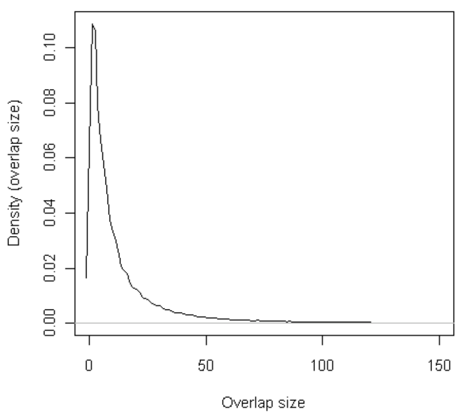

Figure 4. Probability mass function of the overlap size variable (zooming on the non-disjoint pair range).

Additionally, since the McLaughlin's weighting function improves its performance when larger values of n are used, we may conclude that those (few) users that share several items in common are the most important to obtain precise and accurate recommendations. This is because this method would rank higher the users with overlap values larger than n, contrary to the effect of Herlocker's weighting function.

4. CONCLUSIONS AND FUTURE WORK

We have proposed and evaluated three collaborative filtering neighbour selection approaches based on the preference overlap between users. We have computed the user preference overlap using three weighting functions, namely intersection, Herlocker's, and McLaughlin's filtering, and have compared them with other neighbour selection strategies based on user similarity and trust.

We find that overlap-based neighbour selection results in considerable performance improvements with respect to similarity and trust filtering approaches in terms of accuracy and precision. Moreover, the performance results obtained by the proposed approach are particularly positive for small neighbourhoods, enabling computational cost savings, although it suffers a decrease in recommendation coverage. These improvements remain stable for different values of the parameter required by our approach – an overlap threshold n, which controls how similar the Herlocker's and McLaughling's functions are with respect to the intersection size.

In the future we shall explore other user similarity functions besides the Pearson's correlation, which was the only one explored in this paper. Furthermore, we plan to exploit further recommendation input dimensions other than ratings – such as user interaction logs – in order to study if the user preference overlap-based filtering is also helpful when ratings are not explicitly available.

5. ACKNOWLEDGMENTS

This work was supported by the Spanish Government (TIN2011-28538-C02-01) and the Government of Madrid (S2009TIC-1542).

6. REFERENCES

[1] Baeza-Yates R. and Ribeiro-Neto B. 2011. *Modern Information Retrieval: The Concepts and Technology behind Search*. 2nd ed. Addison-Wesley Professional.

[2] Desrosiers C. and Karypis G. 2011. A Comprehensive Survey of Neighborhood-based Recommendation Methods. In: *Recommender Systems Handbook*. Springer-Verlag, Berlin, Heidelberg, 107–144.

[3] Golbeck J. 2009. Trust and nuanced profile similarity in online social networks. *ACM Trans Web.* 3, 4, 1–33.

[4] Herlocker J., Konstan J.A., and Riedl J. 2002. An Empirical Analysis of Design Choices in Neighborhood-Based Collaborative Filtering Algorithms. *Information Retrieval* 5, 4, 287–310.

[5] Hwang C.S. and Chen Y.P. 2007. Using Trust in Collaborative Filtering Recommendation. In *Proceedings of the 20th international conference on Industrial, engineering, and other applications of applied intelligent systems*. Springer-Verlag, Berlin, Heidelberg, 1052–1060.

[6] Jin R., Chai J.Y., and Si L. 2004. An automatic weighting scheme for collaborative filtering. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in Information Retrieval*. ACM Press, New York, NY, 337–344.

[7] Koren Y. 2008. Factorization meets the neighborhood: a multifaceted collaborative filtering model. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM Press, New York, NY, 426–434.

[8] McLaughlin M.R. and Herlocker J.L. 2004. A collaborative filtering algorithm and evaluation metric that accurately model the user experience. In *Proceedings of the 27th annual international ACM SIGIR conference on Research and development in Information Retrieval*. ACM Press, New York, NY, 329–336.

[9] O'Donovan J. and Smyth B. 2005. Trust in recommender systems. In *Proceedings of the 10th international conference on Intelligent User Interfaces*. ACM Press, New York, NY, 167–174.

[10] Resnick P., Iacovou N., Suchak M., Bergstorm P., and Riedl J. 1994. GroupLens: An Open Architecture for Collaborative Filtering of Netnews. In *Proceedings of the 1994 Conference on Computer Supported Cooperative Work*. ACM Press, New York, NY, 175–186.

[11] Said A., Jain B.J., and Albayrak S. 2012. Analyzing weighting schemes in collaborative filtering: cold start, post cold start and power users. In *Proceedings of the 27th Annual ACM Symposium on Applied Computing*. ACM Press, New York, NY, 2035–2040.