

Addressing the New User Problem with a Personality Based User Similarity Measure

Marko Tkalčič, Matevž Kunaver, Andrej Košir, and Jurij Tasič

University of Ljubljana Faculty of electrical engineering,
Tržaška 25, 1000 Ljubljana, Slovenia

{marko.tkalcic,matevz.kunaver,andrej.kosir,jurij.tasic}@fe.uni-lj.si
<http://ldos.fe.uni-lj.si>

Abstract. The new user problem is a recurring problem in memory based collaborative recommender systems (MBCR). It occurs when a new user is added to the system and there are not enough information to make a good selection of the user's neighbours. As a consequence, the recommended items have poor correlation with the user's interests. We addressed the new user problem by observing the user similarity measure (USM). In this paper we present two novelties that address the new user problem : (i) the usage of a personality based USM to alleviate the new user problem and (ii) a method for establishing the boundary of the cold start period. We successfully used a personality based USM that yielded significantly better recommender performance in the period where the new user problem occurs. Furthermore we presented a new methodology for assessing the boundary of the period where the new user problem occurs.

Keywords: memory based collaborative recommender system, new user problem, personality based user similarity measure

1 Introduction

The new user problem is an important issue in memory based collaborative recommender systems [Adomavicius and Tuzhilin, 2005]. It occurs when a new user joins the system and there are no (or there are too few) overlapping ratings to calculate good estimates of user similarities with rating-based user similarity measures (USM). We will denote this initial period as the cold start period (CSP). The consequences of being in the CSP are bad rating predictions for unseen items and thus poor quality of the recommender system. Usually, the new user problem (NUP) has been addressed by introducing content-based approaches which resulted in hybrid systems [Adomavicius and Tuzhilin, 2005, Ahn, 2008]. Once the system has enough overlapping items it is not in the CSP and rating based USM can be used.

We introduced a personality-based USM using the five factor model (FFM) in Tkalčič et al. [2009]. The same approach was later used by Hu and Pu [2010] for the NUP in a music recommender system. In this paper we present (i) the

results of the proposed USM in a CF recommender system for images and (ii) a methodology for assessing the boundary of the cold start period.

The proposed approach to use a personality-based USM in the NUP allows us to calculate user similarities immediately, without waiting for the user to rate several items. The underlying assumption for choosing personality as the basis for the proposed user similarity measure is that people with similar personalities have similar tastes for products. In psychology, personality is described as a set of factors that account for the majority of between-user variance in emotive, interpersonal, experiential and attitudinal styles [John and Srivastava, 1999].

The second novelty is a statistical method for determining at which point the new user problem stops occurring. A review of literature showed that authors either (i) did not set limits for the CSP [Schein et al., 2002] or (ii) provided limits without further argumentation, e.g. Massa and Bhattacharjee [2004] defined *cold start users as the users who have expressed less than 5 ratings*. We propose to determine the boundary of the CSP with a statistical approach, as the number of ratings where the recommender’s performance stops being significantly lower than the performance with higher number of ratings given by the user.

2 The new user problem

The new user problem in collaborative filtering recommenders is described as the period from the moment when a user joins the system to the moment when there are enough ratings to yield a stable list of neighbours (i.e. users with similar preferences). We rewrote this description from various sources [Adomavicius and Tuzhilin, 2005, Schein et al., 2002, Ahn, 2008]. To the best of the authors’ knowledge no formal definition of the new user problem period is available.

In this section we define the boundary of the CSP. Let us have a user u joining the system. The user starts using the system and gives ratings $r(u, h)$ to items $h \in H$ where $H \subset \{h_1 \dots h_J\}$, a set of J items. At any given moment the user has given n ratings to n different items which yields the set

$$R_u^n = \{r_1^u \dots r_n^u\} \quad (1)$$

The boundary of the new user problem period (the CSP) for the selected user is the number of ratings N_u^{CS} after which the system starts to yield stable sets of users. The consequence of a stable set of users is a stable confusion matrix of recommended items. We define that the confusion matrix is stable if a sequence of F -measure values, has statistically equivalent means at different n .

We choose the F measure as a scalar measure of the confusion matrix. We denote the F measure when n ratings have been used to calculate neighbours as F^n . We define the CSP boundary as the point N where the means of F values of the sets

$$R_u^{NJ} = \{F^N \dots F^J\} \quad R_u^{(N-1)J} = \{F^{(N-1)} \dots F^J\} \quad (2)$$

are significantly different.

In Tkalčič et al. [2009] we presented a user similarity measure that takes two vectors $\mathbf{b}_i = (b_{i1} \dots b_{i5})$ and $\mathbf{b}_j = (b_{j1} \dots b_{j5})$ containing the personality values of two users u_i and u_j and yields the scalar similarity value

$$d_W(\mathbf{b}_i, \mathbf{b}_j) = \sqrt{\sum_{l=1}^5 w_l (b_{il} - b_{jl})^2} \quad (3)$$

We use the proposed user similarity measure to calculate similarities in the CSP.

3 Experiment

The goal of the experiment was to (i) assess the CSP boundary and to (ii) see whether the personality-based USM performs better in the CSP. We conducted two experiments with the CF recommender system: (i) one with the personality based USM and (ii) one with the rating based USM.

We used the LDOS-PerAff-1 [Tkalčič et al., 2010] dataset which contained all data necessary to carry out our experiments. The dataset provided the usage history (i.e. the log of users' interactions) of 52 users consuming 70 content items (a subset of images from the IAPS¹ dataset) and giving explicit ratings to each item. The users' task was to assess the images for their computer's wallpaper. The users' personalities vectors \mathbf{b} were assessed using the IPIP50² questionnaire. We used the personality based USM as defined in Eq. 3 to calculate the distances between the users. We calculated the predicted ratings based on the neighbours' ratings using the adjusted Pearson's coefficient as defined in Kunaver et al. [2007]. We then compared the predicted ratings with the ground truth ratings which yielded the confusion matrix.

We calculated the rating based USM $d(u_i, u_j)$ between two arbitrary users u_i and u_j based on their respective ratings $e(u, h)$ of the overlapping items h_m , where m is the index of the items that both the users have rated

$$d(u_i, u_j) = \sqrt{\sum_m (e(u_i, h_m) - e(u_j, h_m))^2} \quad (4)$$

The dataset used in our experiments had a full ratings-items table without missing values with I users and K ratings. To simulate the CSP we determined a usage history path in the form of a random sequence of ratings, for each user separately. We iterated through cold start stages s from one (the user has given only one rating) to K (the user has rated all items) for each user separately. At each stage $1 \leq s \leq K$ we performed the recommender procedure and calculated the confusion matrix for the observed user u at the observed stage s . We chose the F measure as the performance measure of the recommender system. The

¹ <http://csea.phhp.ufl.edu/media.html>

² http://ipip.ori.org/New_IPIP-50-item-scale.htm

experimental procedure thus yielded a table of F values at different stages $s \in \{1..K\}$ and for each user $u \in U$.

3.1 Evaluation methodology

We compared the performance of the rating based USM and the personality based USM by testing the hypothesis $H_0 : \mu_R = \mu_{B5}$ at different cold start stages using the t-test. The value μ_R represents the mean F values using the rating based USM and μ_{B5} represents the mean F values using the personality based USM.

We determined the position of the CSP boundary by testing the hypothesis $H_0 : \mu_s = \mu_{s-K}$ where μ_s represents the mean F value at stage s and μ_{s-K} represents the mean F value from stages $s+1$ to K , where K is the last observed cold start stage.

4 Results

When seeking for the CSP boundary we calculated the p values which are shown in Fig. 1. On the dataset used we observed that $p < 0.05$ occurs when the cold start stage is $s < 6$.

We analyzed the CSP by graphing the quality rate of the recommender(F) versus the number of ratings used. At each cold start stage s the F measures for each user were calculated.

The results of the t test showed that, on the dataset used, the personality based USM yields a significantly higher mean of F values than the rating based USM when the number of ratings taken in account for the calculation of the neighbours is lower than 50 (see Fig. 2). When the number of ratings is higher than 50 the means of F values for both similarity measures are not significantly different at $\alpha = 0.05$.

5 Discussion and conclusion

Experimental results showed that the personality based USM performs significantly better than the rating based USM in cold start conditions. A positive outcome is also the fact that the personality USM is statistically equivalent to the rating based USM which makes it a good candidate for a complete replacement of the rating based USM.

However, the results presented here were verified only on the specific dataset and we don't have any ground to conclude that the presented approach is useful also in other domains. We do speculate that hedonistic-content domains would benefit from the presented approach but this should be verified as future work.

The main drawback of the personality based USM is the difficulty of acquisition of end users' personality parameters. There are two main obstacles in this: (i) it is annoying for the end user to fill in questionnaires and (ii) the acquisition

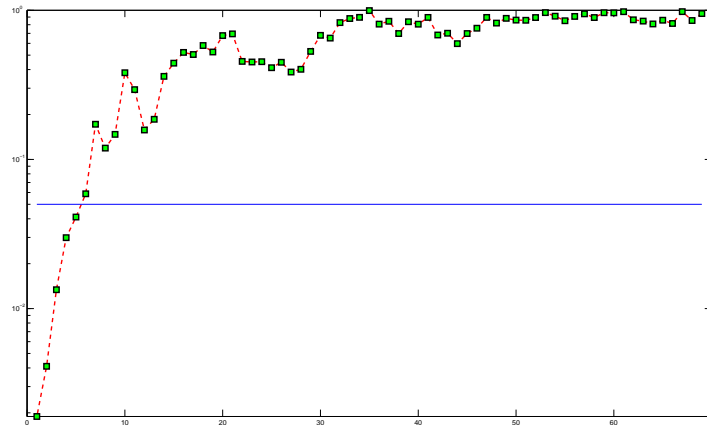


Fig. 1. p values of the t test for the CSP boundary. On the dataset used the CSP occurs when $s < 6$.

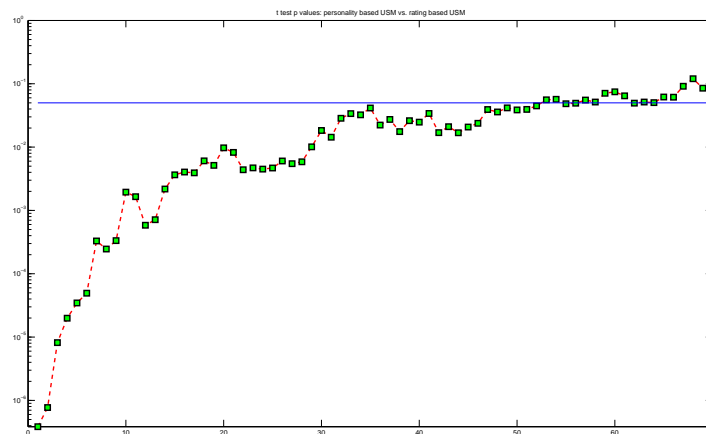


Fig. 2. p values of the t test of the comparison of the personality based USM and rating based USM.

of personality data raises ethical and privacy issues that need to be addressed first. The progress beyond the state of the art here is the knowledge that personality does account for between-users variance in entertainment applications.

In the lack of existing methodologies for assessing the boundaries of the new user problem we chose a statistical approach. We acknowledge that further investigations should be conducted to determine how to test for the CSP boundary and that these investigations might conclude that a different approach is more suitable.

We provided a methodology for the assessment of the new user boundary. The results presented should not be taken for granted and several repetitions of the procedure should be carried out on different datasets.

In this paper we have evaluated a personality based USM under cold start conditions. The results showed that the personality based USM performed significantly better than the rating based USM. Furthermore we described a methodology for the assessment of the CSP border. Both novelties are important in the field of memory based collaborative filtering recommender systems and should be further explored.

References

- G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6):734–749, 2005.
- H.J. Ahn. A new similarity measure for collaborative filtering to alleviate the new user cold-starting problem. *Information Sciences*, 178(1):37–51, 2008.
- R. Hu and P. Pu. Using Personality Information in Collaborative Filtering for New Users. *Recommender Systems and the Social Web*, page 17, 2010.
- Oliver P. John and Sanjay Srivastava. The big five trait taxonomy: History, measurement, and theoretical perspectives. In Lawrence A. Pervin and Oliver P. John, editors, *Handbook of Personality: Theory and Research*, pages 102–138. Guilford Press, New York, second edition, 1999. URL <http://www.uoregon.edu/~sanjay/pubs/bigfive.pdf>.
- Matevž Kunaver, Tomaž Požrl, Matevž Pogačnik, and Jurij Tasič. Optimisation of combined collaborative recommended systems. *International Journal of Electronic Communications*, 61:433–443, 2007.
- P. Massa and B. Bhattacharjee. Using Trust in Recommender Systems: An Experimental Analysis. In *Trust management: second international conference, iTrust 2004, Oxford, UK, March 29-April 1, 2004: proceedings*, page 221. Springer-Verlag New York Inc, 2004.
- A.I. Schein, A. Popescul, L.H. Ungar, and D.M. Pennock. Methods and metrics for cold-start recommendations. In *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 253–260. ACM, 2002.
- Marko Tkalčič, Matevž Kunaver, Jurij Tasič, and Andrej Košir. Personality based user similarity measure for a collaborative recommender system. In C. Peter, E. Crane, L. Axelrod, H. Agius, S. Afzal, and M. Balaam, editors, *Proceedings of the 5th Workshop on Emotion in Human-Computer Interaction - Real world challenges*, pages 30–37. Fraunhofer Verlag, September 2009.
- Marko Tkalčič, Jurij Tasič, and Andrej Košir. The LDOS-PerAff-1 Corpus of Face Video Clips with Affective and Personality Metadata. In Michael Kipp, editor, *Proceedings of the LREC 2010 Workshop on Multimodal Corpora: Advances in Capturing, Coding and Analyzing Multimodality*, 2010.