Personality-Aware Collaborative Filtering: An Empirical Study in Multiple Domains with Facebook Data

Ignacio Fernández-Tobías and Iván Cantador

Universidad Autónoma de Madrid, 28049 Madrid, Spain
{ignacio.fernandezt,ivan.cantador}@uam.es

Abstract. In this paper we investigate the incorporation of information about the users' personality into a number of collaborative filtering methods, aiming to address situations of user preference scarcity. Through empirical experiments on a multi-domain dataset obtained from Facebook, we show that the proposed personality-aware collaborative filtering methods effectively –and consistently in the studied domains– increase recommendation performance, in terms of both precision and recall. We also present an analysis of relationships existing between user preferences and personality for the different domains, considering the users' gender and age.

Keywords: recommender systems, collaborative filtering, personality, user similarity.

1 Introduction

Collaborative filtering has been shown to be one of the most successful approaches for providing personalized recommendations. It has, however, limitations when there is a scarcity of user preferences, which leads to a restricted coverage of the user-item preference space (sparsity problem) and a weak capability of recommending items to a new user (cold start problem).

Aiming to address this situation, approaches have been proposed to exploit additional information about users and items, increasing the density of the user-item preference space or capturing user similarities that do not (merely) depend on personal preferences. Among the existing input signals that have been exploited to establish those similarities, psychological aspects –such as emotions, moods and personality– have been gaining increasing attention by the recommender systems community [12].

Personality, as defined in psychology, is an organized and dynamic set of characteristics possessed by a person that uniquely influence his or her cognitions, motivations, emotions and behaviors in certain situations. In particular, recent studies have shown that personality –accounting for user response patterns– correlates with user preferences in different domains, such as music [14, 16], movies and TV shows [3, 13], books and magazines [2, 15], and websites [10]. These correlations are the basis with which traits and types of personality may help explain why people prefer one option to other, and could be used to improve personalization services.

M. Hepp and Y. Hoffner (Eds.): EC-Web 2014, LNBIP 188, pp. 125–137, 2014.

In fact, several authors [9, 11, 17, 18] have already explored how user personality can be exploited to improve collaborative filtering recommendations. However, due to the current difficulty of obtaining information about both user preferences and personality, proposed approaches have been evaluated with relatively small datasets in single domains. Addressing these limitations, in this paper we present an empirical study comparing state of art and novel personality-aware collaborative filtering methods with a large multi-domain dataset obtained from Facebook. We show that the proposed methods effectively –and consistently in the studied domains– increase recommendation performance, in terms of both precision and recall. Moreover, differently to previous work, we present an analysis of relationships between user preferences and personality in several domains, considering the users' gender and age. We show that these demographic attributes lead to different user preference-personality patterns, which calls for further work in the personality-aware recommendation research agenda.

The reminder of the paper is organized as follows. Section 2 reviews related work. Section 3 presents the evaluated personality-aware collaborative filtering methods, and Section 4 reports and discusses obtained experimental results. Complementing these results, Section 5 analyzes relationships between user preferences and personality in the studied domains. Finally, Section 6 ends with conclusions and future research lines.

2 Related Work

2.1 Personality Modeling

Different models have been proposed to characterize and represent human personality. Among them, the Five Factor (FF) model [4] is considered one of the most comprehensive, and has been the mostly used to build user personality profiles [9]. The FF model establishes five broad domains or dimensions –called factors and commonly known as the Big Five– to describe human personality: *openness*, *conscientiousness*, *extraversion*, *agreeableness* and *neuroticism*.

The openness (OPE) factor reflects a person's tendency to intellectual curiosity, creativity and preference for novelty and variety of experiences. The conscientiousness (COS) factor reflects a person's tendency to show self-discipline and aim for personal achievements, and have an organized (not spontaneous) and dependable behavior. The extraversion (EXT) factor reflects a person's tendency to seek stimulation in the company of others, and put energy in finding positive emotions. The agreeableness (AGR) factor reflects a person's tendency to be kind, concerned, truthful and cooperative towards others. Finally, the neuroticism (NEU) factor reflects a person's tendency to experience unpleasant emotions, and refers to the degree of emotional stability and impulse control.

The measurement of the FFs comprises items that are self-descriptive sentences or adjectives, commonly presented in the form of short tests. In this context, the International Personality Item Pool[1] (IPIP) is a publicly available collection of items for use in psychometric tests, and the 20-100 item IPIP proxy for Costa and McCrae's

[1] International Personality Item Pool (IPIP), `http://ipip.ori.org`

NEO-PI-R test [5] is one of the most popular and widely accepted questionnaires to measure the Big Five in adult men and women without overt psychopathology.

In this paper we use the FF model to represent user personality. In particular, we utilize the dataset provided by myPersonality project[2] [1], which has personality and entertainment preference profiles of Facebook users. The dataset contains the FF scores of 3.1 million users, collected by using 20 to 336 item IPIP questionnaires. With this dataset we aim to exploit the relationships between user preferences and personality for recommendation.

2.2 Relationships between User Preferences and Personality

Personality influences how people make their decisions [12]. In particular, it has been shown that people with similar personality characteristics are likely to have similar preferences. In [16] Rentfrow and Gosling investigated how music preferences are related with personality in terms of the FF model, showing that “reflective” people with high openness usually have preferences for jazz, blues and classical music, and “energetic” people with high degree of extraversion and agreeableness usually appreciate rap, hip-hop, funk and electronic music. Also in the music domain and using the FF model, Rawlings and Ciancarelli [14] observed that openness and extraversion are the personality factors that best explain the variance in personal preferences. They showed that people with high openness tend to like diverse music styles, and people with high extraversion are likely to have preferences for popular music.

In the movie domain, Chausson [3] presented a study showing that people open to experiences are likely to prefer comedy and fantasy movies, conscientious individuals are more inclined to enjoy action movies, and neurotic people tend to like romantic movies. Afterwards, Odi et al. [13] explored the relationships between personality factors and induced emotions in movies for different social contexts –e.g. watching a movie alone or with someone else–, and observed different patterns in experienced emotions as functions of the extraversion, agreeableness and neuroticism factors.

Extending the spectrum of analyzed domains, in [15] Rentfrow et al. investigated the relations between personality factors and user preferences in several entertainment domains, namely movies, TV shows, books, magazines and music. They focused their study on five personality-based content categories: aesthetic, cerebral, communal, dark and thrilling. The authors observed positive and negative relationships between such categories and some of the personality factors, e.g., they showed that aesthetic contents relate positively with agreeableness and negatively with neuroticism, and that cerebral contents correlate with extraversion. Also considering several domains – movies, TV shows, books and music–, in [2] we presented a preliminary study on the relations between personality types and entertainment preferences. Analyzing a large dataset of personality factor and domain genre preference user profiles, we extracted personality-based user stereotypes for each genre, and inferred association rules and similarities between types of personality of people with preferences for particular genres. Finally, in the multi-domain scenario of the Web, Kosinski et al. [10]

[2] myPersonality project, `http://mypersonality.org`

presented a study revealing meaningful psychologically relationships between user preferences and personality for certain websites and website categories.

Continuing our previous work [2], in this paper we analyze the relationships between user preferences and personality in multiple entertainment domains. Here we extend the number of considered domain genres and, differently from other studies, conduct the analysis by clustering the user profiles according to demographic attributes, namely gender and age. We show that there are differences in the correlations between user preferences and personality depending on such attributes.

2.3 Personality-Aware Recommender Systems

Once extracted, personality factors could be used to build personality-based user profiles that may be exploited by personalized information retrieval and filtering approaches [8, 11]. In [18] Tkalčič et al. applied and evaluated three user similarity metrics for the heuristic-based collaborative filtering strategy: a typical rating-based similarity, a similarity based on the Euclidean distance with FF data, and a similarity based on a weighted Euclidean distance with the FF data. Their reported results showed that approaches using FF data perform statistically equivalent or better than the rating-based approach, especially in cold-start situations. In her PhD dissertation [11], Nunes explored the use of personality user profile composed of NEO-IPIP items and facets in addition to the Big Five factors, showing that fine-grained personality user profiles help achieve better recommendations. Following the findings of Rentfrow and Gosling [16], in [9] Hu and Pu presented a collaborative filtering approach based on the correlations between personality types and music preferences, in which the similarity between two users is estimated by means of the Pearson's correlation coefficient on the users' FF scores. Combining this approach with a rating-based collaborative filtering, the authors showed significant improvements over the baseline of considering only ratings data. In [17] Roshchina proposed an approach that extracts FF-based profiles by analyzing hotel reviews written by users, and incorporates these profiles into a nearest neighbor algorithm to enhance personalized recommendations. Finally, in [6] Elahi et al. presented an active learning technique that exploits user personality to accurately acquire user ratings for collaborative filtering in cold-start situations.

Instead of using the Euclidean distance as done in [18], in this paper we propose to use the cosine-based distance to establish (dis)similarities between users in the collaborative filtering strategy for both preference- and personality-based user profiles. Similarly to [11], we explore alternative representations of personality user profiles, but focusing on Big Five data instead of considering the large number of NEO-IPIP items and facets. For comparison purposes, we implement and evaluate the Pearson-based user similarity presented in [9], but we also assess a user similarity based on Spearman's correlation, aiming to capture non-linear relationships between user personality factors. Finally, as done in [17], we evaluate collaborative filtering strategies based on nearest neighbors, but differently to that work, in this paper we report recommendation performance results for several neighborhood sizes, and evaluate hybrid approaches that combine user preference and personality data.

3 Personality-Aware Collaborative Filtering Methods

In the recommender systems literature, the preference of a user $u \in \mathcal{U}$ for certain item $i \in \mathcal{I}$ is commonly represented as a rating $r_{u,i} \in \mathcal{R}$, where $\mathcal{R}$ is a totally ordered set, e.g. non negative integers or real numbers within a certain range. For each user u, a recommender system aims to predict ratings $r_{u,i}$ of items i unrated by u, and suggest the items with the highest predicted ratings $\tilde{r}_{u,i}$.

In particular, heuristic-based collaborative filtering strategies compute $\tilde{r}_{u,i}$ as an aggregate of the ratings of some other (usually, the most similar) users v. More formally, $\tilde{r}_{u,i} = \mathrm{aggr}_{v \in \mathcal{N}_u} r_{v,i}$, where $\mathcal{N}_u$ denotes the set of N users who are the most similar to u, and is usually referred as the set of neighbors or neighborhood. Several aggregation functions aggr have been proposed in the literature. One of the most widely used, and that we adopt in this paper (for unary/binary ratings) is:

$$\tilde{r}_{u,i} = \overline{r_u} + \kappa \sum_{v \in \mathcal{N}_u} sim(u,v) \cdot (r_{v,i} - \overline{r_v})$$

where $sim(u,v)$ is a function that measures the similarity between two users u and v, $\overline{r_u}$ is the average value of user u'ratings, and κ is a normalization factor, which is usually set as $\kappa = 1/\sum_{v \in \mathcal{N}_u} sim(u,v)$.

Together with the rating aggregate function, the user similarity function is a central component that characterizes a collaborative filtering strategy. A standard similarity metric is based on the Pearson's correlation coefficient:

$$sim(u,v) = \frac{\sum_{i \in \mathcal{I}_u \cap \mathcal{I}_v} (r_{u,i} - \overline{r_u})(r_{v,i} - \overline{r_v})}{\sqrt{\sum_{i \in \mathcal{I}_u \cap \mathcal{I}_v} (r_{u,i} - \overline{r_u})^2 \sum_{i \in \mathcal{I}_u \cap \mathcal{I}_v} (r_{v,i} - \overline{r_v})^2}}$$

where $\mathcal{I}_u \subseteq \mathcal{I}$ is the set of items rated by user u. This metric is not applicable when the ratings are either unary (such as the Facebook's likes) or binary (such as the YouTube's thumps up/down). In these cases, a user similarity based on the Jaccard's coefficient is commonly used:

$$sim_{pref}(u,v) = \frac{|\mathcal{I}_u \cap \mathcal{I}_v|}{|\mathcal{I}_u \cup \mathcal{I}_v|} \tag{1}$$

This is in fact the similarity metric we consider as baseline in our experiments. We shall refer the collaborative filtering strategy that utilizes $sim_{pref}(u,v)$, without considering the users' personality, as CF. Analogously to [9], in order to incorporate personality information into the recommendation process, we shall study a hybrid recommendation strategy that linearly combines $sim_{pref}(u,v)$ with a personality-based user similarity $sim_{pers}(u,v)$ as follows:

$$sim(u,v) = \lambda \cdot sim_{pref}(u,v) + (1-\lambda) \cdot sim_{pers}(u,v) \tag{2}$$

where $\lambda \in [0,1]$ controls the influence of user preferences and personality on the recommendation process. For λ values close to 1, user preferences are more relevant, while for λ values close to 0, personality gets higher relevance. In the subsequent sections, we propose several formulations of $sim_{pers}(u,v)$ that yield the personality-aware collaborative filtering methods we empirically compare.

Let $p_u = \{p_{u,ope}, p_{u,cos}, p_{u,ext}, p_{u,agr}, p_{u,neu}\}$ be the user u's personality-based profile composed of u's Big Five scores. Previous approaches have considered the components $p_{u,k}$ ranging in a numeric interval, e.g. [1,5] and [1,100], as derived from the IPIP questionnaires. Here, we consider this representation as well, and refer it as *continuous* personality profile. Nonetheless, aiming to explore simpler personality-aware user profiles and similarities, we also propose to study a personality representation in which each component $p_{u,k}$ takes one of a limited set of categorical values –namely low, medium, and high–, obtained from the corresponding original Big Five scores. We refer this representation as *discrete* personality profile.

3.1 Cosine-Based Personality User Similarity Methods (COS)

The first personality-based user similarity we evaluate is the cosine-based similarity:

$$sim_{pers\_cos}(u,v) = \frac{\sum_k p_{u,k} \cdot p_{v,k}}{\sqrt{\sum_k p_{u,k}^2}\sqrt{\sum_k p_{v,k}^2}} \tag{3}$$

Its incorporation in formula (2) produces a recommendation method that we call COS. Depending on whether continuous or discrete personality profiles are used, we shall refer to COS-c and COS-d methods, respectively. These methods will have different implementations depending on the λ value used, i.e., COS-c-λ and COS-d-λ, with $0 < \lambda < 1$. The case of $\lambda = 1$ (exploiting only user preferences) is equivalent to the CF method, and the case of $\lambda = 0$ will be referred as COS-c-pers and COS-d-pers, since only user personality information is exploited.

3.2 Pearson-Based Personality User Similarity Methods (PEA)

The second personality-based user similarity we evaluate was proposed by Hu and Pu in [9], and is based on the Pearson's correlation coefficient:

$$sim_{pers\_pea}(u,v) = \frac{\sum_k (p_{u,k} - \overline{p_u})(p_{v,k} - \overline{p_v})}{\sqrt{\sum_k (p_{u,k} - \overline{p_u})^2 \sum_k (p_{v,k} - \overline{p_v})^2}} \tag{4}$$

Its integration in (2) produces a recommendation method that we call PEA. Similarly to the COS case, for the PEA similarity, we shall evaluate the PEA-c-per, PEA-d-per, PEA-c-λ and PEA-d-λ methods, with $0 < \lambda < 1$.

3.3 Spearman-Based Personality User Similarity Methods (SPE)

The third and last personality-based user similarity we propose is based on the Spearman's correlation coefficient:

$$sim_{pers\_spe}(u,v) = \frac{\sum_k (s_{u,k} - \overline{s_u})(s_{v,k} - \overline{s_v})}{\sqrt{\sum_k (s_{u,k} - \overline{s_u})^2 \sum_k (s_{v,k} - \overline{s_v})^2}} \tag{5}$$

where $s_{u,k}$ is the position of $p_{u,k}$ in the decreasing order ranking of u's Big Five scores. Analogously to previous cases, the incorporation of this similarity into (2) produces a recommendation method that we call SPE, and which will be instantiated and evaluated in the SPE-c-per, SPE-d-per, SPE-c-λ and SPE-d-λ methods.

4 Experiments

4.1 Dataset

The dataset used in our experiments is part of the database made publicly available in myPersonality project [1]. myPersonality is a Facebook application with which users take psychometric tests and receive feedback on their scores. The users allow the application to record personal information from their Facebook profiles, such a demographic and geo-location data, *likes*, status updates, and friendship relations, among others. In particular, as of March 2014, the tool has let record a database with 46 million Facebook likes of 220,000 users for 5.5 million items of diverse nature – people (actors, musicians, politicians, sportsmen, writers, etc.), objects (movies, TV shows, songs, books, games, etc.), organizations, events, etc.–, and the Big Five scores of 3.1 million users, collected using 20 to 336 item IPIP questionnaires.

Due to the size and complexity of the database, in this paper we restrict our study to a subset of its items. Specifically, we selected all likes (ratings) associated to items belonging to one of the following 3 categories: Movie genre, Book genre, and Musical genre. Thus, for instance, selected items belonging to the Movie genre category are movie genres such as comedy, action, adventure, drama, and science fiction. Note that we do not take into account a large number of potential valuable items, such as particular movies preferred by users. Next, we selected those users of the dataset that had ratings for the considered items. Once the items and users were selected, we conducted text processing operations to consolidate morphological derivations of certain item names (e.g. science fiction, science-fiction, sci-fi, and sf). Table 1 shows some statistics about the number of users, items and ratings in the considered domains. The minimum, maximum, and average (standard deviation) numbers of ratings per user were: 1, 19 and 1.73 (1.14) for movies, 1, 15 and 1.57 (1.08) for books, and 1, 61 and 3.68 (3.66) for music. In the table, users are grouped by gender and by age according to Erikson's psychosocial stages [7] (see Table 2).

Table 1. Statistics of the used dataset (gender and age of some users were not declared)

Domain	#users						
	all	***female***	***male***	***adolescent***	***young adult***	***middle-aged adult***	***advanced-aged adult***
movies	16168	9827	6341	2833	3086	1610	189
books	15251	9919	5332	2672	2916	1577	202
music	17980	10924	7056	3164	3467	1898	234
Domain (#items)	**#likes (sparsity)**						
	all	***female***	***male***	***adolescent***	***young adult***	***middle-aged adult***	***advanced-aged adult***
movies (268)	27921 (99.36%)	17073 (99.35%)	10848 (99.36%)	4850 (99.36%)	5192 (99.37%)	2643 (99.39%)	289 (99.43%)
books (305)	23882 (99.49%)	15717 (99.48%)	8165 (99.50%)	4105 (99.50%)	4478 (99.50%)	2493 (99.48%)	336 (99.45%)
music (1175)	66079 (99.69%)	38898 (99.70%)	27181 (99.67%)	11628 (99.69%)	12055 (99.70%)	5876 (99.74%)	654 (99.76%)

As explained in Section 3, we propose to evaluate the recommendation methods utilizing both continuous and discrete user personality profiles. Specifically, for building the discrete profiles, we transform each original personality factor score in [1, 5] into one of the following categories: "low", "medium", and "high." Considering the distribution of personality factor scores of all users in the dataset, for each personality factor, the *low* category is assigned to those scores that are below the 33th percentile, the *medium* category is assigned to those scores that are between the 33th and 66th percentiles, and the *high* category is assigned to those scores that are between the 66th and 100th percentiles. Table 2 shows the obtained intervals.

Table 2. Considered categories for user ages (left) and personality factors (right)

Age category	Age interval
adolescent	< 20 years old
young adult	[20, 39]
middle-aged adult	[40, 50]
advanced-aged adult	> 50 years old

Personality factor	Personality factor categories and intervals		
	low	***medium***	***high***
OPE	[1.00, 3.75)	[3.75, 4.25)	[4.25, 5.00]
COS	[1.00, 3.05)	[3.05, 3.75)	[3.75, 5.00]
EXT	[1.00, 3.25)	[3.25, 4.00)	[4.00, 5.00]
AGR	[1.00, 3.25)	[3.25, 4.00)	[4.00, 5.00]
NEU	[1.00, 2.50)	[2.50, 3.10)	[3.10, 5.00]

4.2 Results

In our experiments we empirically compared 61 methods: CF, COS-c-λ, COS-d-λ, PEA-c-λ, PEA-d-λ, SPE-c-λ, and SPE-d-λ, with $\lambda = 0, 0.1, 0.2, \dots, 0.9$. The methods were evaluated in the three considered domains, by means of their precision, recall and F-measure values for the top *k* recommendations –P@*k*, R@*k* and F@*k*– with *k* = 1,…,5, averaged by 5-times 5-fold cross validation. All methods were executed with neighborhood sizes of 5, 10, 15, 20 and 25. Tables 3, 4 and 5 show the methods' configurations with best results (in terms of average F-measure values). All differences with CF's values are statistically significant (2-tailed Wilcoxon, p≤0.05).

Table 3. Avg. P@*k*, R@*k* and F@*k* values achieved by the best performing methods in the <u>movies</u> domain. Values higher than CF's are in bold, and the highest values are underlined.

Method	Neighbors	P@1	P@2	P@3	P@4	P@5	R@1	R@2	R@3	R@4	R@5	F@1	F@2	F@3	F@4	F@5
CF	20	.146	.382	.350	<u>.480</u>	.490	.025	.140	.247	.268	.298	.043	.205	.290	.344	.371
COS-c-pers	5	.099	.242	.290	.339	.353	**.099**	**.241**	**.284**	**.292**	.294	.099	.241	.287	.314	.321
COS-c-0.3	15	**.154**	**.400**	**<u>.488</u>**	.463	.460	.024	.119	.242	**.306**	.291	.042	.183	**.324**	**.368**	.356
COS-d-pers	10	.068	.228	.282	.252	.298	**.068**	**.228**	**.282**	.252	.298	**.068**	**.228**	.282	.252	.298
COS-d-0.1	10	**.149**	**.405**	**.433**	.456	**.532**	**.062**	**.213**	**.306**	**.276**	**.313**	**.088**	**.279**	**.359**	.345	**.394**
PEA-c-pers	5	.095	.267	.337	.375	.373	**.095**	**<u>.267</u>**	**.333**	**.343**	**.323**	**.095**	**.267**	**.335**	**.358**	.346
PEA-c-0.2	10	**<u>.178</u>**	**<u>.442</u>**	**.466**	.439	.372	**.071**	**.236**	**.350**	**.340**	**.314**	**.102**	**<u>.308</u>**	**<u>.400</u>**	**.383**	.341
PEA-d-pers	5	.061	.210	.302	.280	.321	**.061**	**.209**	**.294**	.267	**.315**	**.061**	**.209**	**.298**	.273	.318
PEA-d-0.1	10	**.149**	**.400**	**.436**	.476	**<u>.540</u>**	**.075**	**.221**	**.311**	**.284**	**<u>.317</u>**	**.100**	**.285**	**.363**	**.356**	**<u>.399</u>**
SPE-c-pers	5	.102	.243	**.351**	.357	.261	**<u>.102</u>**	**.240**	**.331**	**.316**	.223	**.102**	**.241**	**.341**	.335	.241
SPE-c-0.1	10	**.162**	.380	**.429**	.463	.300	**.089**	**.257**	**.322**	**<u>.347</u>**	.227	**<u>.115</u>**	**<u>.307</u>**	**.368**	**<u>.397</u>**	.258
SPE-d-pers	5	.057	.226	.299	.296	.330	**.050**	**.200**	**<u>.266</u>**	.237	.254	**.053**	**.212**	.282	**.263**	.287
SPE-d-0.2	10	.139	**.406**	**.455**	.463	.410	**.047**	**.188**	**<u>.297</u>**	**.288**	.261	**.070**	**.257**	**.359**	**.355**	.319

Table 4. Avg. P@k, R@k and F@k values achieved by the best performing methods in the books domain. Values higher than CF's are in bold, and the highest values are underlined.

Method	Neighbors	P@1	P@2	P@3	P@4	P@5	R@1	R@2	R@3	R@4	R@5	F@1	F@2	F@3	F@4	F@5
CF	15	.097	.146	.189	.227	.291	.026	.081	.129	.199	.245	.041	.104	.153	.212	.266
COS-c-pers	5	.052	.102	.152	.201	.243	**.052**	**.102**	**.150**	.187	.196	**.052**	.102	.151	.194	.217
COS-c-0.1	15	.095	**.150**	**.235**	**.292**	.270	**.027**	**.081**	**.147**	**.254**	.219	**.042**	**.105**	**.181**	**.272**	.242
COS-d-pers	10	.035	.098	.163	.236	.286	**.035**	**.098**	**.163**	**.236**	**.286**	.035	.098	**.163**	**.236**	**.286**
COS-d-0.3	10	.087	**.153**	**.236**	**.266**	**.342**	**.029**	**.085**	**.164**	**.205**	**.276**	**.044**	**.109**	**.194**	**.232**	**.305**
PEA-c-pers	10	.050	.088	.130	.208	.266	**.050**	**.088**	**.130**	**.208**	**.266**	**.050**	.088	.130	.208	**.266**
PEA-c-0.2	10	.087	**.159**	**.196**	**.263**	.277	**.041**	**.103**	**.150**	**.216**	.242	**.056**	**.125**	**.170**	**.237**	.258
PEA-d-pers	10	.043	.100	.164	.231	.256	**.043**	**.100**	**.164**	**.231**	**.256**	**.043**	.100	**.164**	**.231**	.256
PEA-d-0.5	10	.077	**.164**	**.221**	**.248**	**.381**	.023	**.090**	**.144**	**.209**	**.298**	.035	**.116**	**.174**	**.227**	**.334**
SPE-c-pers	5	.040	.105	.170	.164	.218	**.040**	**.104**	**.169**	.158	.197	.040	**.104**	**.169**	.161	.207
SPE-c-0.5	20	.092	**.156**	.184	**.229**	.272	**.036**	**.112**	**.164**	**.221**	**.272**	**.052**	**.130**	**.173**	**.225**	**.272**
SPE-d-pers	10	.036	.083	.141	**.242**	.272	**.032**	.076	**.129**	**.218**	**.255**	.034	.079	.135	**.229**	.263
SPE-d-0.3	10	.086	**.157**	**.222**	**.249**	**.355**	**.031**	**.087**	**.152**	.173	**.260**	**.046**	**.112**	**.180**	**.204**	**.300**

Table 5. Avg. P@k, R@k and F@k values achieved by the best performing methods in the music domain. Values higher than CF's are in bold, and the highest values are underlined.

Method	Neighbors	P@1	P@2	P@3	P@4	P@5	R@1	R@2	R@3	R@4	R@5	F@1	F@2	F@3	F@4	F@5
CF	5	.130	.192	.230	.250	.247	.066	.118	.148	.169	.172	.088	.146	.180	.202	.203
COS-c-pers	5	.033	.066	.083	.094	.107	.033	.065	.083	.094	.106	.033	.066	.083	.094	.106
COS-c-0.8	5	**.135**	**.197**	**.238**	.248	**.247**	**.069**	**.120**	**.151**	.167	**.172**	**.092**	**.149**	**.184**	.200	**.203**
COS-d-pers	5	.067	.073	.081	.080	.089	**.067**	.073	.081	.080	.088	.067	.073	.081	.080	.088
COS-d-0.7	5	**.130**	.189	.219	.222	.236	**.070**	**.119**	**.149**	.164	**.182**	**.091**	**.146**	.177	.189	**.206**
PEA-c-pers	5	.032	.058	.073	.090	.099	.032	.058	.073	.090	.098	.032	.058	.073	.090	.098
PEA-c-0.9	5	**.131**	**.194**	**.234**	.244	**.255**	**.067**	.115	**.149**	.167	**.184**	**.089**	.144	**.182**	.198	**.214**
PEA-d-pers	5	.064	.071	.081	.079	.091	.064	.071	.081	.078	.088	.064	.071	.081	.078	.089
PEA-d-0.9	5	**.134**	**.194**	.228	**.250**	**.248**	**.069**	**.119**	.144	**.169**	**.177**	**.091**	**.148**	.177	**.202**	**.206**
SPE-c-pers	5	.024	.060	.081	.096	.101	.024	.059	.081	.095	.099	.024	.059	.081	.095	.100
SPE-c-0.9	5	**.130**	**.193**	.227	.241	.241	**.067**	.115	.145	.163	**.177**	**.088**	.144	.177	.194	**.204**
SPE-d-pers	5	.070	.074	.073	.073	.089	.063	.067	.066	.067	.079	.066	.070	.069	.070	.084
SPE-d-0.9	5	**.133**	**.196**	.224	.245	.240	.063	.108	.133	.155	.162	.086	.139	.167	.190	.193

In general, for the three domains, the proposed hybrid recommendation methods that exploit information about the users' **personality** outperform the baseline CF method. It has to be noted that, probably due to the scarcity of user preferences (see Table 1), those methods that give more weight to user personality (i.e., those methods with λ values close to 0) obtain the highest precision and recall values. This is not the case in the music domain, where the average number of rated items per user is larger than in the movies and books domains (see Section 4.1), and the relationships between user preferences and personality are the weakest ones (as will be shown in Section 5).

In terms of average F-measure values, for most cases in the movies and books domains, the **discrete profiles** let obtained better results than the continuous profiles. Regarding the **user similarities**, there is no consensus on which is the best alternative for the studied domains. While in the music domain the COS methods are the best performing, in the movies and books domains the PEA and SPE methods respectively achieve the highest precision and recall values.

4.3 Discussion

The obtained results show that the proposed 3-category personality profiles let generate better recommendations than those obtained with numeric personality profiles, which are the ones used previously in the literature. The discretization process of the Big Five presented in this paper follows a simple approach based on the personality factor distributions in the whole dataset. Approaches dependent on the target domain, considering a different number of personality factor categories, or based on certain user aspects may help improve generated recommendations.

The nature of the utilized dataset, which comprises *like* preferences for domain genres, entailed the use of the Jaccard user similarity (formula 1) in the heuristic-based collaborative filtering strategy for the top *N* recommendation task. If either binary *like/unlike* records or numeric ratings were available as user preferences, more elaborated collaborative filtering methods –e.g. those based on the Pearson's and Spearman's correlation coefficients– could be investigated for both the top *N* and rating prediction recommendation tasks.

Moreover, although precision and recall metrics are appropriate and widely used for evaluating the top *N* recommendation task, other metrics could be explored when assessing personality-aware collaborative filtering methods. In particular, measuring recommendation diversity may be of special interest. As shown in previous studies [2, 14], people open to experience tend to have preferences for more diverse types of items.

5 Relationships between User Preferences and Personality

As done in previous work (see Section 2.2), here we report relationships existing between user preferences and personality. Tables 6 and 7 show linear correlation values between preference- and personality-based user similarities for all distinct pairs of users in each domain. We report such correlation values for different groups of users based on their gender and age, as explained in Section 4.1. All values marked with asterisk (*) are statistically significant at $p<0.05$.

When considering **all users** in a domain, it can be seen that correlation values are close to 0, meaning that there is no (linear) relations between user preferences and personality factors. If we consider **age-based groups**, the relations are strong for adolescent users in the movies domain, and are moderate for middle-aged adults in the movies and music domains. On the other hand, if we consider **gender-based groups**, we can observe that the correlation values are higher between female users in the movies domain and, in less degree, in the books domain. In the music domain, there is no clear distinction for correlations between male and female users. Finally, we note that there are no clear patterns between the correlations obtained with the cosine- and Pearson/Spearman-based personality profiles.

Table 6. Linear correlations between preference- and personality-based user similarities

Domain	minimum #likes per user	Cosine-based Personality User Profiles											
		all ages			*adolescent*			*young adult*			*middle-aged adult*		
		all	*female*	*male*	*all*	*female*	*male*	*all*	*female*	*male*	*all*	*female*	*male*
movies	4	.039* (792920)	.053* (296846)	.032* (119804)	.143* (23065)	.155* (7562)	.155* (4182)	.043* (21101)	.045* (7223)	.066* (3586)	.034* (5105)	.074* (2177)	-.084* (605)
	5	.067* (137646)	.080* (47735)	.064* (23125)	**.213*** (3364)	**.356*** (933)	.115* (733)	-.027 (3793)	-.082* (1352)	.060 (597)	.059 (1094)	.067 (388)	-.126 (169)
	6	.120* (19408)	.136* (5867)	.128* (3891)	**.267*** (435)	**.545*** (91)	.150 (120)	.095 (605)	.124 (158)	.083 (136)	**.212*** (153)	**.420*** (28)	.007 (45)
books	4	.037* (230644)	.043* (107132)	.043* (24008)	.058* (5911)	.074* (2495)	.097* (746)	.014 (7012)	.013 (2484)	.031 (909)	.080* (2270)	.118* (1149)	.124* (197)
	5	.042* (58495)	.054* (26688)	.017 (6238)	.037 (729)	.043 (271)	-.013 (112)	.077* (2073)	.099* (855)	.113 (252)	.057* (757)	.107* (374)	.050 (65)
	6	.033* (15926)	.062* (7196)	-.058* (1691)	.060 (279)	.060 (144)	-.419 (17)	.010 (489)	.018 (225)	.169 (46)	.048 (110)	.125 (28)	.115 (23)
music	4	.022* (13962828)	.024* (5065849)	.019* (2248669)	.027* (433486)	.022* (146624)	.033* (77412)	.030* (428909)	.034* (150762)	.015* (73019)	-.008* (12311)	.009* (42641)	-.008* (12311)
	8	.040* (1763904)	.048* (597027)	.037* (311186)	.056* (51022)	.084* (17597)	.038* (8726)	.028* (52463)	.003* (15661)	.045* (10877)	.037* (1025)	.115* (2851)	.037* (1025)
	16	.025* (41503)	.011 (9672)	.038* (11090)	.116* (1375)	.125* (344)	.169* (335)	.083* (1113)	-.096 (153)	**.229*** (421)	-.156 (15)	**.464*** (19)	-.156 (15)

Table 7. Linear correlations between preference- and personality-based user similarities

Domain	minimum #likes per user	Spearman-based Personality User Profiles											
		all ages			*adolescent*			*young adult*			*middle-aged adult*		
		all	*female*	*male*	*all*	*female*	*male*	*all*	*female*	*male*	*all*	*female*	*male*
movies	4	.006* (767201)	.001 (288179)	.028* (115290)	.032* (22263)	.013 (7233)	.099* (4061)	.022* (20130)	.010 (7029)	.047* (3326)	.032* (4909)	.046* (2083)	-.106* (593)
	5	.024* (133411)	.012* (46436)	.049* (22387)	.137* (3260)	**.210*** (902)	.098* (706)	.005 (3616)	-.067* (1321)	.134* (548)	.047* (1061)	.074 (377)	-.137 (169)
	6	.063* (18982)	.046* (5754)	.098* (3787)	**.248*** (424)	**.574*** (89)	.094 (112)	.103* (594)	-.063 (157)	.273* (135)	-.008* (149)	.113 (28)	-.025 (45)
books	4	.037* (224332)	.034* (104144)	.050* (23385)	.029* (5766)	.033 (2429)	.058 (731)	.011 (6823)	-.007 (2790)	.027 (881)	.103* (2210)	.119* (1119)	.066 (184)
	5	.042* (56869)	.041* (25946)	.027* (6064)	.053 (728)	.083 (266)	-.019 (111)	.006 (2023)	-.019 (841)	.038 (243)	.134* (734)	.115* (363)	-.001 (64)
	6	.027* (15477)	.039* (6998)	-.015 (1639)	.086 (273)	.101 (140)	-.457 (17)	-.017 (477)	-.089 (222)	.168 (43)	.097 (107)	-.066 (27)	.064 (23)
Music	4	-.001* (13552095)	.007* (4928335)	-.008* (2175265)	-.007* (147674)	-.017* (142927)	.005 (75410)	.001 (415860)	.008* (147105)	-.006 (70182)	.002 (97066)	.013* (41223)	-.012 (12012)
	8	.014* (1712038)	.025* (580695)	.004* (301306)	.015* (48739)	.014 (17132)	.004 (8511)	.028* (50968)	.040* (15304)	.027* (10495)	.056* (7152)	.080* (2803)	.017* (1009)
	16	-.030* (40470)	-.012 (9453)	-.043* (10774)	.046 (1341)	-.004 (337)	.052 (329)	.063* (1095)	.005 (153)	.113* (413)	-.048 (72)	.263 (19)	-.403 (14)

6 Conclusions and Future Work

In this paper we have presented an empirical study comparing several collaborative filtering methods that effectively exploit information about the users' personality. In addition to evaluating state of the art approaches, we have assessed new personality-aware recommendation strategies and user profiles, and have used large datasets in several domains. Moreover, we have analyzed relationships between user preferences and personality considering the users' gender and age, which has revealed differences in the preference-personality patterns, calling for further work in the future.

In our study, however, we have focused on broad representations of the users' personality by assuming the same relevance for all the Big Five personality factors. As previous work has shown [3, 14, 16], personality factors may influence differently on user preferences for certain types of items. In the future we shall investigate how fine grained relationships between user preferences and personality factors can be exploited in recommendation. Moreover, we also envision the consideration of personality facets (in addition to or instead of personality factors) as a potential way to exploit personality information in collaborative filtering.

Our investigation has revealed there are differences in the relationships between user preferences and personality according to the users' gender and age. Nonetheless, we have not exploited such particularities for recommendation purposes yet. In this context, other user attributes, such as educational attainment, could be taken into account as well.

Finally, we also plan to evaluate new personality-aware collaborative filtering approaches, e.g. methods based on factor models, and strategies focused on neighbor selection and rating matrix enrichment, instead of on user similarities.

References

1. Bachrach, Y., Kohli, P., Graepel, T., Stillwell, D.J., Kosinski, M.: Personality and Patterns of Facebook Usage. In: Proceedings of the 4th Web Science Conference (WebSci 2012), pp. 24–32 (2012)
2. Cantador, I., Fernández-Tobías, I., Bellogín, A.: Relating Personality Types with User Preferences in Multiple Entertainment Domains. In: Proc. of the 1st Workshop on Emotions and Personality in Personalized Services, pp. 13–28 (2013)
3. Chausson, O.: Who Watches What? Assessing the Impact of Gender and Personality on Film Preferences. myPersonality project, University of Cambridge (2010)
4. Costa, P.T., McCrae, R.R.: Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) Manual. Psychological Assessment Resources (1992)
5. Goldberg, L.R., Johnson, J.A., Eber, H.W., Hogan, R., Ashton, M.C., Cloninger, C.R., Gough, H.G.: The International Personality Item Pool and the Future of Public-Domain Personality Measures. Journal of Research in Personality 40, 84–96 (2006)
6. Elahi, M., Braunhofer, M., Ricci, F., Tkalcic, M.: Personality-Based Active Learning for Collaborative Filtering Recommender Systems. In: Baldoni, M., Baroglio, C., Boella, G., Micalizio, R. (eds.) AI*IA 2013. LNCS (LNAI), vol. 8249, pp. 360–371. Springer, Heidelberg (2013)
7. Erikson, E.H.: Childhood and Society. W. W. Norton and Company (1950)
8. Hu, R., Pu, P.: A Study on User Perception of Personality-Based Recommender Systems. In: De Bra, P., Kobsa, A., Chin, D. (eds.) UMAP 2010. LNCS, vol. 6075, pp. 291–302. Springer, Heidelberg (2010)
9. Hu, R., Pu, P.: Enhancing Collaborative Filtering Systems with Personality Information. In: Proceedings of the 5th ACM Conference on Recommender Systems, pp. 197–204 (2011)
10. Kosinski, M., Stillwell, D.J., Kohli, P., Bachrach, Y., Graepel, T.: Personality and Website Choice. In: Proceedings of 4th ACM Conference on Web Science, pp. 251–254 (2012)
11. Nunes, M.A.S.N.: Recommender Systems based on Personality Traits: Could Human Psychological Aspects Influence the Computer Decision-making Process? VDM Verlag (2009)

12. Nunes, M.A.S.N., Hu, R.: Personality-based Recommender Systems: An Overview. In: Proceedings of the 6th ACM Conference on Recommender Systems, pp. 5–6 (2012)
13. Odić, A., Tkalčič, M., Tasič, J.F., Košir, A.: Personality and Social Context: Impact on Emotion Induction from Movies. In: Proc. of the 1st Workshop on Emotions and Personality in Personalized Services (2013)
14. Rawlings, D., Ciancarelli, V.: Music Preference and the Five-Factor Model of the NEO Personality Inventory. Psychology of Music 25(2), 120–132 (1997)
15. Rentfrow, P.J., Goldberg, L.R., Zilca, R.: Listening, Watching, and Reading: The Structure and Correlates of Entertainment Preferences. Journal of Personality 79(2), 223–258 (2011)
16. Rentfrow, P.J., Gosling, S.D.: The Do Re Mi's of Everyday Life: The Structure and Personality Correlates of Music Preferences. Journal of Personality and Social Psychology 84(6), 1236–1256 (2003)
17. Roshchina, A.: TWIN: Personality-based Recommender System. MSc thesis, Institute of Technology Tallaght, Dublin (2012)
18. Tkalčič, M., Kunaver, M., Košir, A., Tasič, J.F.: Addressing the New User Problem with a Personality Based User Similarity Measure. In: Proc. of the 2nd Workshop on User Models for Motivational Systems (2011)