Personality Based User Similarity Measure for a Collaborative Recommender System

Marko Tkalčič
University of Ljubljana Faculty of electrical engineering
Tržaška 25
Ljubljana, Slovenia
marko.tkalcic@fe.uni-lj.si

Matevž Kunaver
University of Ljubljana Faculty of electrical engineering
Tržaška 25
Ljubljana, Slovenia
matevz.kunaver@fe.uni-lj.si

Jurij Tasič
University of Ljubljana Faculty of electrical engineering
Tržaška 25
Ljubljana, Slovenia
jurij.tasic@fe.uni-lj.si

Andrej Košir
University of Ljubljana Faculty of electrical engineering
Tržaška 25
Ljubljana, Slovenia
andrej.kosir@fe.uni-lj.si

ABSTRACT

We propose a novel approach for calculating the user similarity for collaborative filtering recommender systems that is based on the big five personality model. Experimental results showed that the performance of the proposed measures is either equal or better (depending on the measure under evaluation) than the ratings based measures used in state-of-the-art collaborative recommender systems. This makes the proposed approach, with its benefits in terms of computational complexity, for calculating user similarities a very promising one for future collaborative recommender systems that will be more affect-oriented.

Categories and Subject Descriptors

H.1.2 [**Models and Principles**]: User Machine Systems; H.3.3 [**Information Storage and Retrieval**]: Information Search and Retrieval

Keywords

collaborative filtering, big five personality model, personality based user similarity measure, affective computing

1. INTRODUCTION

Recommender systems are gaining on importance as the amount of available multimedia content is expanding rapidly. The main drivers are digital television, video on demand and web2.0 services like flickr and youtube. It is very hard for end users to find suitable content in large databases where most of the content is irrelevant. This is why recommender systems that filter only relevant items for each user are important. Based on how recommendations are made Adomavicius and Tuzhilin [2005] differ between content based recommenders (CBR), collaborative filtering recommenders (CF) and hybrid recommenders (HR). CBR systems are based on user's inclinations towards specific attributes of the content item. The knowledge about the user is stored in data structures called user models. CF systems recommend items that similar users have liked in the past. HR use both CBR and CF approaches. The CF recommenders are further divided into memory-based and model-based recommenders. While memory-based recommenders work with the whole database of users' past ratings and preferences the model-based recommenders compile the whole database into a smaller data structure, the model [Pennock et al., 2000, Adomavicius and Tuzhilin, 2005].

In this paper we focus on memory-based CF systems. CF recommenders are based on the presumption that when the similarity between two users is high both users will like similar items. The similarity measure is thus a crucial part of any CF system.

1.1 Related Work and Problem Statement

State of the art multimedia recommender solutions (both CF and CBR) work on a very non-affective, technical basis as can be deducted from the overviews of Adomavicius and Tuzhilin [2005] and Lew et al. [2006]. CBR systems exploit metadata fields (e.g. genre) to build user and item profiles upon which the recommendations are made. On the other hand, CF recommenders rely solely on past user ratings to build the recommendations.

The strict technical approach in recommender systems that ignores the users' affective experiences during the consumption of multimedia content is odd because the entertainment industry is based on *giving people emotions*. In fact, Masthoff and Gatt [2006] treat the user's satisfaction when consuming multimedia content as an affective state. While there have been some research efforts in CBR systems to exploit emotive responses of users during content consumption for producing better recommendations [González et al., 2004, Lekakos and Giaglis, 2006, Shan et al., 2009, Tkalčič et al., 2009] CF recommenders have not received any such attention. The problem that we are addressing is the clear lack of affective elements in state of the art CF recommender systems.

The main reason probably lies in the fact that, by their nature, CF recommenders ignore content and user metadata where affective information could be stored. They are based on the presumption that *close* users (user *closeness* is calculated using a user similarity measure) have similar preferences for multimedia content. So, if users u_1 and u_2 are close and user u_1 has liked the content item h it is highly probable that user u_2 is going to like the item h as well. Thus the only place where one could make good use of emotions in CF systems is the user similarity measure.

One possible solution would be to construct a user similarity measure based on users' past emotive responses on same items. This would require to have an automatic emotion detection system. Such algorithms do exist [Donato et al., 1999, Picard and Daily, 2005, Zeng et al., 2009] but there is another problem: the similarities between users would need to be calculated on a regularly basis which is computationally very expensive. Each time a user consumes an item new information are added to the usage history which is the basis for the calculation of the similarity. This implies that after each update of the usage history the similarities between users should be recalculated.

1.2 Proposed Solution

We propose to use a similarity measure that yields, for each user u, a list of close neighbours that have in common a similar emotive response pattern to content items. We suggest to exploit end users' personalities to build such a similarity measure. According to McCrae and John [1992] personality tries to explain the individual differences in emotive reactions to common stimuli. So personality does affect the emotive response of users during multimedia content consumption. The big five personality model [McCrae and John, 1992, John and Srivastava, 1999] appears to be a promising instrument since the user's personality is described by a tuple of five numerical values which are convenient for computer calculations. Especially the extraversion and neuroticism factors seem to be tightly connected with individuals' emotive responses [Yik et al., 2002]. Despite the appearance that personality and emotions are two distinct theoretical and empirical streams substantial evidence has been found that they are tightly connected [Carver et al., 2000, Luminet et al., 2000, Davidson, 2001]. The main advantage of the proposed solution is that personality doesn't change with time. Thus once we have the big five tuple for a user we can calculate the neighbours in advance and have the list ready at any time which is a big improvement in terms of computational complexity.

In addition to the introduction of elements of affective computing in CF recommenders, the proposed solution has three advantages over standard memory based CF recommenders: (i) it solves the new user problem by introducing an initial questionnaire, (ii) it has lower computational requirements for the calculation of similarities between users and (iii) it lowers the impact of the sparsity problem as the calculation of similarities does not depend on ratings. The main drawback is that the initial questionnaire could be annoying for users.

Our hypothesis is that the CF recommender system's performance is not significantly lower when using the personality based measure than using the standard rating based measures.

1.3 Paper Outline and Notations

We provide the argumentation how personality affects emotions in section 2. In section 3 we give a description of the CF recommender. The experimental procedure along with the evaluated similarity measures is described in section 4. The results of the experiment are given in section 5. These results are further discussed in section 6.

The notations used throughout this paper are given in table 1.

2. PERSONALITY IMPLICATIONS ON EMOTIONS

Westen [1999] states that personality refers to the enduring patterns of thought, feeling, motivation and behaviour that are expressed in different circumstances. According to McCrae and John [1992] the big five factor model of personality is a hierarchical organization of personality traits in terms of five basic dimensions: extraversion (E), agreeableness (A), conscientiousness (C), neuroticism (N) and openness (O). These are the most important ways in which the individuals differ in their enduring emotional, interpersonal, experiential, attitudinal and motivational styles. The order of the factors is important as the first two (E and A) account for the largest percentage of variance in personality [John and Srivastava, 1999]. Each of the factors encompasses more specific traits.

Yik et al. [2002] reported that all five factors influence feelings and emotional behaviour. This is especially evident for the E and N superfactors while there is a nonzero correlation between the other three factors (O, C and A) and affect behaviour.

Johnson [2009] provided a description of the five factors and their subfactors. In the following of this section we give an overview based on Johnson [2009].

The E factor tells the degree of engagement with the external world (in case of high values) or the lack of it (low values). The subfactors of E are friendliness, gregariousness, assertiveness, activity level, excitement-seeking and cheerfulness. Extrovert people (high score on the E factor) tend to react with enthusiasm and often have positive emotions while introverted people (low score on the E factor) tend to be quiet, low-key and disengaged in social interactions.

The N factor refers to the tendency of experiencing negative feelings. People with high N values are emotionally reactive. They tend to respond emotionally to relatively neutral

Sign	Description
U	set of users
H	set of items
$u \in U$	single user
$h \in H$	single item
$e_L(u,h) \in \Omega_L$	Likert rating given to item h by user u
Ω_L	set of discrete Likert values (1 to 5)
$e(u,h) \in \Omega$	binary rating given to item h by user u
$\Omega = \{C_0, C_1\}$	set of binary ratings/classes (C_1 =relevant / C_0 =non-relevant)
$d(u_i, u_j) \in [0, \infty)$	distance function between users u_i and u_j
$sim(u_i, u_j) = \frac{1}{1+d} \in (0, 1]$	similarity measure between users u_i and u_j
$\vec{b}_i = (b_{i1}, \ldots, b_{i5})$	vector of big five values for the user u_i
$F_j(\vec{b}) = \lambda_j \vec{b}^{\top}.\vec{v}_j$	principal components of the big five values
$d_R(u_i, u_j)$	rating based distance function between users u_i and u_j
$d_E(\vec{b}_i, \vec{b}_j)$	big five euclidian distance function between users' u_i and u_j respective big five vectors $\vec{b}_i$ and $\vec{b}_j$
$d_{PCA}(\vec{b}_i, \vec{b}_j)$	weighted big five euclidian distance function between users' u_i and u_j respective big five vectors $\vec{b}_i$ and $\vec{b}_j$. The weights are the PCA components
$\bar{e}_L^{NN}(u,h)$	average rating of user's u nearest neighbours for the item h in the Ω domain
$\bar{e}_L^{P}$(h)	average rating of all users for the item h in the Ω_L domain
$\hat{e}_L(u,h) \in [1, 5]$	estimated numerical rating of item h for user u
$\hat{e}(u,h) \in \Omega$	estimated binary rating of item h for user u
k	number of neighbours
m	threshold
n	number of neighbours who have rated item h
$H_R(u)$	set of items relevant for the user u

Table 1: Notations used throughout the article

stimuli. They are often in a bad mood which strongly affects their thinking and decision making. Low N scorers are calm, emotionally stable and free from persistent bad mood. The subfactors are anxiety, anger, depression, self-consciousness, immoderation and vulnerability.

The distinction between imaginative, creative people and down-to-earth, conventional people is described by the O factor. High O scorers are typically individualistic, non conforming and are very aware of their feelings. They can easily think in abstraction. People with low O values tend to have common interests. They prefer simple and straightforward thinking over complex, ambiguous and subtle. The subfactors are imagination, artistic interest, emotionality, adventurousness, intellect and liberalism.

The C factor concerns the way in which we control, regulate and direct our impulses. People with high C values tend to be prudent while those with low values tend to be impulsive. The subfactors are self-efficacy, orderliness, dutifulness, achievement-striving, self-discipline and cautiousness.

The subdomains of the A factor are trust, morality, altruism, cooperation, modesty and sympathy. The A factor reflects individual differences in concern with cooperation and social harmony.

3. CONTENT FILTERING SCENARIO

Let us have a set of I users $U = \{u_1, u_2, \ldots u_I\}$ that are using our CF recommender system. We also have a set of J multimedia items $H = \{h_1, h_2, \ldots h_J\}$. Each time a user u consumes an item h she/he is required to give an explicit rating to the item which we denote as $e_L(u,h)$. The rating values e_L are taken from a set of discrete Likert values Ω_L which ranges from 1 to 5. The Likert values are mapped to binary ratings $e(u,h)$ from the set of possible values (classes)

$$\Omega = \{C_0, C_1\} \tag{1}$$

where $C_0 =' non-relevant'$ and $C_1 =' relevant'$ by applying the mapping $\Omega_L \to \Omega$:

$$e(u,h) = C_0 : e_L(u,h) <= 3 \tag{2}$$

$$e(u,h) = C_1 : e_L(u,h) > 3. \tag{3}$$

The usage history of a memory based CF system can be represented as a table of all item ratings given by users until the observed moment (see Tab. 2 for a hypothetical example).

	h_1	h_2	...	h_J
u_1	4	2		1
u_2		2		
u_3	2			1
u_4		3		
u_5	3			
u_6	1			2
u_7		5		3
...				
u_I		2		4

Table 2: Usage and rating history table. As all users have not rated all items there are several empty entries.

The user u then chooses an item h to consume and then gives it a rating which is stored in the usage history table.

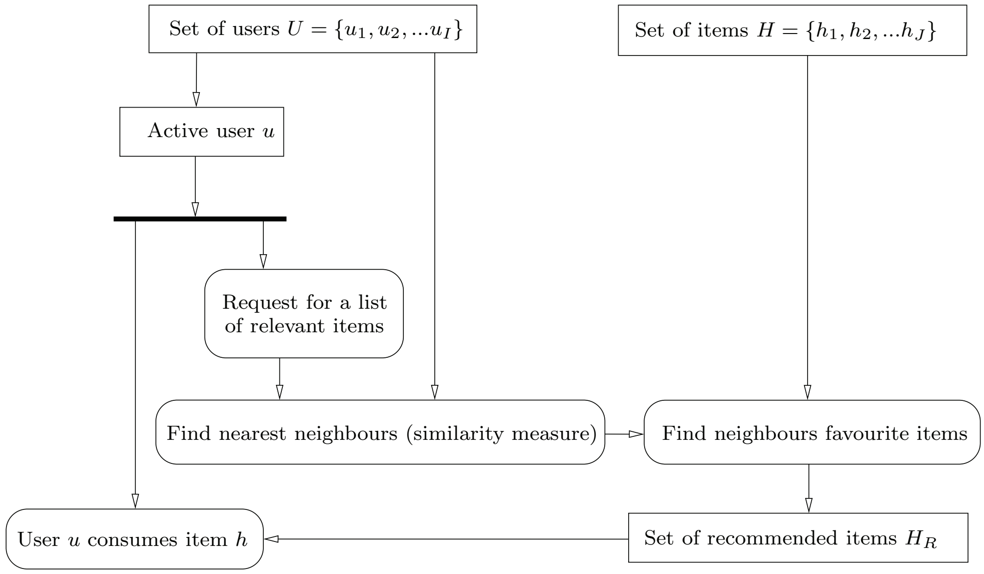

Figure 1: CF recommender system scenario.

3.1 Rating Prediction in Content Filtering Recommenders

The recommending procedure is centred around a single user u (see Fig. 2). When user u requests the CF a list of recommended items the system calculates binary rating predictions $\hat{e}(u, h)$ from the set Ω for all the items that have not been viewed by the user. The list of recommended items $H_R(u) \subset H$ is composed by all items that fall in the C_1 class (relevant items),

$$H_R(u) = \{h : \hat{e}(u, h) = C_1\}. \tag{4}$$

The procedure starts by calculating the list of k nearest neighbours for the user u. The similarity between user u and all other users is calculated using a user similarity measure $sim(u, u_i)$ where $u_i \in U \backslash \{u\}$.

The k users with the highest values $sim(u, u_i)$ are chosen as the k nearest neighbours.

For each item h two ratings that take values from Ω_L are calculated: the user's u nearest neighbours' average rating $\bar{e}_L^{NN}(h, u)$ and the overall average rating $\bar{e}_L^P(h)$. Both average ratings are aggregated into a single numerical rating prediction value $\hat{e}_L$ using the true Bayesian estimate [imd, 2009]

$$\hat{e}_L(u, h) = \frac{n}{n+m}\bar{e}_L^{NN}(u, h) + \frac{m}{n+m}\bar{e}_L^P(h) \tag{5}$$

where n represents the number of neighbours who have rated the item h while m represents the threshold value. If we set m to a lower value then $\hat{e}_L(u, h)$ is more dependent of $\bar{e}_L^{NN}(u, h)$ and vice versa, if we set m to a value close to n then $\hat{e}_L(u, h)$ is more dependent on $\bar{e}_L^P(u, h)$.

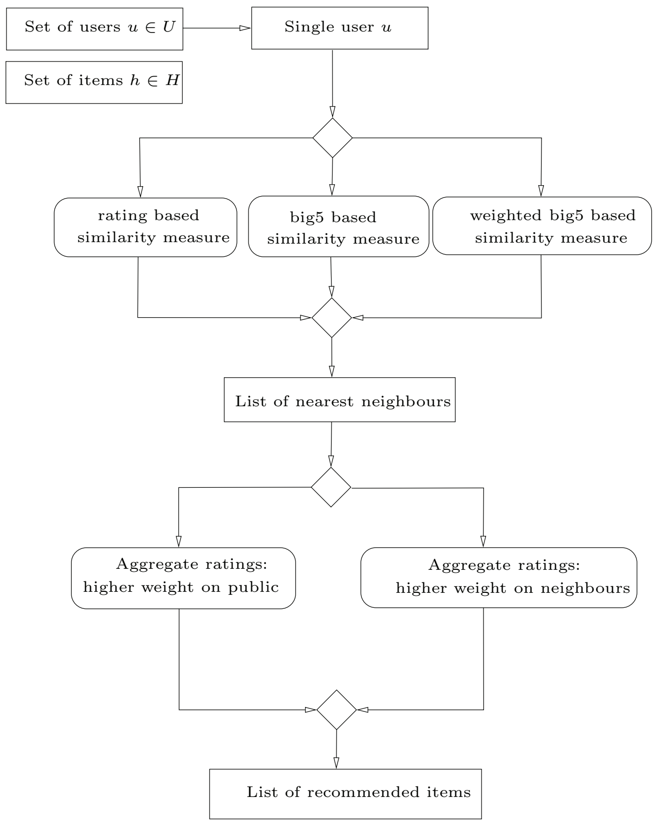

Figure 2: Rating prediction procedure for a selected user u

The estimated rating $\hat{e}_L$ is a numerical value with $\hat{e}_L \in [0, 5]$. The final estimation of the item's binary rating is calculated by thresholding $\hat{e}_L$

$$\hat{e}(u,h) = C_0 : \hat{e}_L(u,h) <= 3 \quad (6)$$

$$\hat{e}(u,h) = C_1 : \hat{e}_L(u,h) > 3 \quad (7)$$

After this procedure has gone through all non viewed items and has thus calculated all the rating predictions the CF recommender compiles a personalized list of recommended items according to equation (4) which is a reduced and thus manageable subset of relevant items $H_R(u)$ among all available items H.

4. MATERIALS AND METHODS

We performed a two stage experiment. First we acquired the dataset needed and then we used it in the CF recommender system. We used the precision P, recall R and F measure to assess the performance of the recommender system with the evaluated similarity measures and thus to validate (or reject) our hypothesis.

4.1 Dataset Acquisition

The subjects involved in the dataset acquisition phase were $N_U = 52$ (15 males) with average age of 18.3 years (SD = 0.56). Each subject was shown a sequence of 70 images taken from the IAPS database [Lang et al., 2005]. The images were carefully chosen to equally cover the widest possible area in the value-arousal space of induced emotions. After viewing each image the subjects were required to rate the image on a Likert scale from 1 (meaning *I don't like it at all*) to 5 (meaning *I like it very much*) in order to advance to the next image. This procedure yielded a full usage history matrix (same structure as table 2 but without missing values).

In order to assess the big five personality values each subject was required to fill in a big five questionnaire. We used the IPIP questionnaire with 50 items [ipi, 2009]. Each of the five factors were covered by 10 items in the questionnaire. The subjects had to describe how accurately each statement describes her/him on a scale from 1 (meaning *very inaccurate*) through 3 (*neither inaccurate nor accurate*) to 5 (*very accurate*). Half of the statements had a positive relation to the describing factor and the other half had a negative relation. Each answer was corrected according to the relation to the factor and distinctly summed and normalized in order to yield separate sums for each factor. Table 3 shows an excerpt of our dataset.

4.2 Content Filtering Recommender Implementation

The acquired dataset was used in the CF recommender system developed by Kunaver et al. [2007].

We evaluated the CF using three different user similarity measures: (i) a standard, rating based measure (see equation (8)), (ii) an Euclidian big five based measure (see equation (9)) and a (iii) weighted Euclidian big five based measure (see equation (10)). After calculating the k nearest neighbours using these measures the item rating prediction (for all items in the dataset) was calculated by combining the average ratings of the neighbours and the average rating of all users (which we denote as *public*). We aggregated both average ratings into the final item rating estimation using the equation (5) with two weight configurations: (i) by putting more weight on the neighbours' average rating $\bar{e}_L^{NN}$ and (ii) by putting more weight on all users' average rating $\bar{e}_L^{P}$ (public). This yielded six CF recommender experiment runs:

- Rating based measure with more weight on neighbours
- Rating based measure with more weight on public
- Euclidian big five based measure with more weight on neighbours
- Euclidian big five based measure with more weight on public
- Weighted Euclidian big five based measure with more weight on neighbours
- Weighted Euclidian big five based measure with more weight on public

We chose $k = 5$ for the number of neighbours which is a rough 10% of all users in our dataset.

The three similarity measures $SIM = \{sim_R, sim_E, sim_{PCA}\}$ were calculated using three different distance measures $D = \{d_R, d_E, d_{PCA}\}$ respectively. For the respective similarity measures $sim \in SIM$ and distance measures $d \in D$ we set $sim = \frac{1}{1+d}$.

We calculated the rating based distance d_R between two users u_i and u_j based on past ratings of both users to all items except the observed one $h_k \in H \backslash \{h\}$

$$d_R(u_i,u_j)^2 = \sum_k (e_L(u_i,h_k) - e_L(u_j,h_k))^2 \quad (8)$$

and the big five distance measures using the users' respective big five vectors $\vec{b}_i$ and $\vec{b}_j$

$$d_E(\vec{b}_i,\vec{b}_j)^2 = \sum_l |b_{il} - b_{jl}|^2 \quad (9)$$

$$d_{PCA}(\vec{b}_i,\vec{b}_j)^2 = \sum_l |F_l(\vec{b}_i) - F_l(\vec{b}_j)|^2 \quad (10)$$

where the weights $F_l(\vec{b}) = \lambda_l \vec{b}^\top . \vec{v}_l$ are the result of the principal component analysis which yields

$$d_{PCA}(\vec{b}_i,\vec{b}_j)^2 = \sum_l |\lambda_l (\vec{b}_i - \vec{b}_j)^\top . \vec{v}_l|^2 \quad (11)$$

A study from Kunaver et al. [2007] showed that the best performance of a CF recommender in terms of F-measure is yielded when one of the two average ratings, $\bar{e}_L^{NN}(h)$ and $\bar{e}_L^{P}(h)$, has a higher weight than the other. In terms of equation (5) we had the parameter n set at $n = 5$ (because of the full dataset all neighbours have rated all items) and we evaluated two m parameter values $m \in \{1, 4\}$.

	BIG5 values					Content ratings					
	E	A	C	N	O	h_1	h_2	h_3	...	h_{J-1}	h_J
u_1	3,2	2,7	2,9	3,5	2,9	4	3	1	...	1	3
u_2	2,1	3,5	3,1	3,4	3,6	2	4	4	...	4	1
u_3	3,2	3,0	2,8	3,2	3,1	2	3	2	...	1	3
...	...	...	...	...	...	...	...	...	...	...	...
u_I	3,3	3,0	3,4	3,9	3,2	4	3	4	...	2	3

Table 3: The acquired dataset

4.3 Evaluation Methodology

Recommender systems are usually assessed based on how well they distinguish items that are relevant for a specific user from those that are non-relevant (see Herlocker et al. [2004]). The performance of this binary classification problem is described with the confusion matrix of correctly and incorrectly classified instances. Focusing on the relevant items a classifier yields four groups: (i) true positives (TP) are items that are relevant to the user and have been correctly classified as relevant, (ii) true negatives (TN) are items that have been correctly classified as non relevant, (iii) false positives (FP) are items that are non-relevant but have been misclassified as relevant and (iv) false negatives (FN) are items that are relevant but have been misclassified as non relevant. Following Herlocker et al. [2004] we calculated three numeric values that describe each a certain aspect of the CF recommender's performance: precision, recall and F-measure.

Precision P is the rate of truly relevant items among all the items classified as relevant by the CF system

$$P = \frac{TP}{TP + FP} \tag{12}$$

Recall R is the rate of returned relevant items to all relevant items

$$R = \frac{TP}{TP + FN} \tag{13}$$

The F measure combines precision and recall in a single numerical value

$$F = \frac{2PR}{P + R} \tag{14}$$

We calculated the P, R and F values for each single user $u \in U$ over all items in the dataset H and for each of the six CF combinations.

In order to prove (or reject) the hypothesis set in Sec. 1.2 we performed the one way analysis of variance to determine whether the differences of mean values of the F measure for the six CF combinations are significantly different or not.

5. RESULTS

In terms of mean values of P, R and F the big five based approaches performed better than the ratings based approaches. The mean values of P, R and F are reported in Tab. 4. The ANOVA analysis (with the significance level $p < 0.05$, we do not report the respective p values here) of the F-measure further showed that all the big five based approaches perform significantly better than the distance based measure with higher weight on the neighbours and are statistically equivalent to the distance based measure with higher weight on the public. The box plot of the values of the F-measure is shown in Fig. 3.

similarity measure	aggregation mode	P	R	F
rating based	neighbours	0.6666	0.5895	0.6268
rating based	public	0.7042	0.7401	0.7232
big5	neighbours	0.6309	0.8533	0.7062
big5	public	0.7093	0.8068	0.7442
weighted big5	neighbours	0.6455	0.8398	0.7165
weighted big5	public	0.7104	0.8064	0.7450

Table 4: Mean values of P, R and F for the different combinations of measures and weighting.

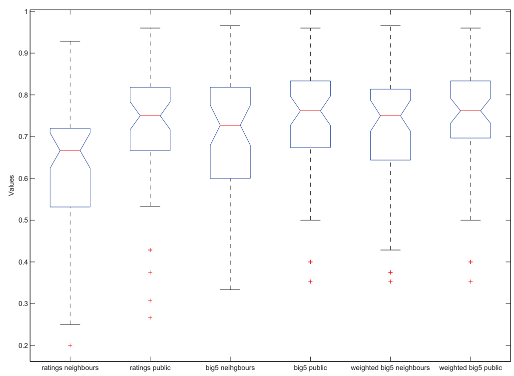

Figure 3: Box plots of the F measure results for six combinations of measures and weighting.

6. DISCUSSION AND FUTURE WORK

We performed an offline experiment of a memory based CF recommender system that relies on end users' personality parameters to determine the nearest neighbours, which is a crucial part of the recommending procedure. We compared four personality based similarity measures and two rating based similarity measures. The CF recommender system's performance results showed that the personality based measures were statistically equivalent or superior (the mean value of F was significantly higher) to the rating based measure which is used in most state-of-the-art CF recommender systems. The proposed approach has several advantages compared to the rating based similarity measures: (i) it is less computationally intensive because the similarities need to be calculated only when a user joins the system while in rating based similarity measures they need to be updated each time a user adds a new rating, (ii) according to the analysis of variance it is statistically equivalent to state-of-the-art rating based methods and (iii) it includes personality and consequently affect in the domain of CF recommenders which are very user oriented. This makes the proposed approach an excellent candidate for more efficient and more affective oriented CF recommender systems.

Unfortunately the proposed system still requires explicit user feedback in the form of ratings which can be annoying for end users. The fact that we introduced a user similarity measure that does not rely on ratings is not enough to give up the explicit feedback. The proposed measure allows the system to find users that have common multimedia interest which is reflected by their common emotive responses based on their personality values. But we still need to know which items are relevant for specific groups of users and this is the reason why we still need to have explicit ratings. Automatic methods for the detection of users' satisfaction (through the analysis of users' facial actions, gestures etc.) would be a valid alternative.

We must also be aware of the fact that personality is not the only parameter that influences our affective responses to stimuli. According to Westen [1999] the human behaviour (and thus emotional dynamics) is better described by *if-then* patterns where situational context plays an equal part as does personality. So in addition to the personality parameters we would also need to know the context (the *if* of the *if-then* pattern) in which the user is during the consumption of the item to improve the performance of the recommender system.

We acknowledge that the biggest drawback of the proposed approach is the need for an initial questionnaire to determine the big five values for each user. Such questionnaires are usually annoying and turn away potential users from using the system. For future work we propose to change the way the questionnaire is implemented into a more friendly and funny thing that can work as an attractor. A similar approach, called photo profiling, has been taken by Berger et al. [2007] that implemented a game where each user (of a personalized tourism destination recommender) had a set of images and she/he had to drag each image into a bin (like it / don't like it). In this way the system calculates two parameters that describe the potential tourist: the pack and the kick factors.

There are several other issues to address in the near future: (i) search for new measures that are based on personality/affect, (ii) include context awareness in the recommender system and (iii) perform more exhausting testings.

References

Imdb top 250. `http://www.imdb.com/chart/top`, June 2009.

International personality item pool: A scientific collaboratory for the development of advanced measures of personality traits and other individual differences. `http://ipip.ori.org/`, June 2009.

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. URL `http://ieeexplore.ieee.org/xpls/abs\_all.jsp?arnumber=1423975`.

Helmut Berger, Michaela Denk, Michael Dittenbach, Andreas Pesenhofer, and Dieter Merkl. Photo-based user profiling for tourism recommender systems. In Giuseppe Psaila and RolandWagner, editors, *E-Commerce and Web Technologies*, volume 4655/2007, pages 46–55. Springer Berlin / Heidelberg, 2007.

C.S. Carver, S.K. Sutton, and M.F. Scheier. Action, emotion, and personality: Emerging conceptual integration. *PERSONALITY AND SOCIAL PSYCHOLOGY BULLETIN*, 26(6):741–751, JUN 2000. ISSN 0146-1672.

R.J. Davidson. Toward a biology of personality and emotion. In Damasio, AR and Harrington, A and Kagan, J and McEwen, BS and Moss, H and Shaikh, R, editor, *UNITY OF KNOWLEDGE: THE CONVERGENCE OF NATURAL AND HUMAN SCIENCE*, volume 935 of *ANNALS OF THE NEW YORK ACADEMY OF SCIENCES*, pages 191–207. NEW YORK ACAD SCIENCES, 2 EAST 63RD ST, NEW YORK, NY 10021 USA, 2001. ISBN 1-57331-310-6.

Gianluca Donato, MArian Stewart Bartlett, Joseph C. Hager, and Paul Ekmanand Terrence J. Sejnowski. Classifying facial actions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 21(10):974–989, October 1999.

G. González, B. López, and de la Rosa J. LL. Managing emotions in smart user models for recommender systems. In *Proceedings of 6th International Conference on Enterprise Information Systems ICEIS 2004*, volume 5, pages 187–194, April 2004.

Jonathan L. Herlocker, Joseph A. Konnstan, Loren G. Terveen, and John T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Transactionson Information Systems*, 22(1), January 2004.

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`.

John A. Johnson. Descriptions used in ipip-neo narrative report. `http://www.personal.psu.edu/faculty/j/5/j5j/IPIPNEOdescriptions.html`, June 2009.

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.J. Lang, M.M. Bradley., and B.N. Cuthbert. International affective picture system (iaps): Affective ratings of pictures and instruction manual. technical report a-6. Technical report, University of Florida, Gainesville, FL, 2005.

George Lekakos and George M. Giaglis. Improving the prediction accuracy of recommendation algorithms: Approaches anchored on human factors. *Interacting with computers*, 18:410–431, 2006.

Michael S. Lew, Nicu Sebe, Chabane Djeraba, and Ramesh Jain. Content-based multimedia information rtrieval: State of the art and challenges. *ACM Transactions on multimedia computing*, 2(1):1–19, February 2006.

O Luminet, E Zech, B Rime, and H Wagner. Predicting cognitive and social consequences of emotional episodes: The contribution of emotional intensity, the Five Factor Model, and alexithymia. *JOURNAL OF RESEARCH IN PERSONALITY*, 34(4):471–497, DEC 2000. ISSN 0092-6566. doi: {10.1006/jrpe.2000.2286}.

Judith Masthoff and Albert Gatt. In pursuit of satisfaction and the prevention of embarrassment: affective state in group recommender systems. *User Modeling and User-Adapted Interaction*, 16(3-4):281–319, September 2006.

Robert R. McCrae and Oliver P. John. An introduction to the five-factor model and its applications. *Journal of Personality*, 60(2):p175 – 215, 06 1992. ISSN 00223506.

David Pennock, Eric Horvitz, Steve Lawrence, and Lee C. Giles. Collaborative filtering by personality diagnosis: A hybrid memory- and model-based approach. In *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence, UAI 2000*, pages 473–480, Stanford, CA, 2000. URL `http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.38.6498`.

Rosalind Picard and Shaundra Briant Daily. Evaluating affective interactions: Alternatives to asking what users feel. In *CHI Workshop on Evaluating Affective Interfaces: Innovative Approaches*, Portland, OR, April 2005.

Man-Kwan Shan, Fang-Fei Kuo, Meng-Fen Chiang, and Suh-Yin Lee. Emotion-based music recommendation by affinity discovery from film music. *Expert Syst. Appl.*, 36(4):7666–7674, 2009. ISSN 0957-4174. doi: http://dx.doi.org/10.1016/j.eswa.2008.09.042.

Marko Tkalčič, Jurij Tasič, and Andrej Košir. Usage of affective parameters in a content-based multimedia recommender system. *Submitted to the Elsevier International Journal of Human-Computer Studies*, 2009.

Drew Westen. *Psychology: Mind, Brain and Culture.* John Wiley and Sons, Inc, 2nd edition, 1999.

Michelle Yik, James A. Russell, Chang-kyu Ahn, Jose Miguel Fernandez Dols, and Naoto Suzuki. Relating the five-factor model of personality to a circumplex model of affect : a five-language study. In R. R. McCrae and J. Allik, editors, *The Five-Factor Model of Personality Across Cultures*, pages 79 – 104. Kluwer Academic Publishers, New York, 2002.

Zhihong Zeng, Maja Pantic, Glenn I. Roisman, and Thomas S. Huang. A survey of affect recognition methods: Audio, visual, and spontaneous expressions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 31(1):39–58, Jan. 2009. ISSN 0162-8828. doi: 10.1109/TPAMI.2008.52.