Available online at www.sciencedirect.com
SCIENCE DIRECT
Interacting with Computers 18 (2006) 410–431
Interacting with Computers
www.elsevier.com/locate/intcom

Improving the prediction accuracy of recommendation algorithms: Approaches anchored on human factors

George Lekakos *, George M. Giaglis [1]

Department of Management Science and Technology, Athens University of Economics and Business, 47 Evelpidon and Lefkados Str., 11362 Athens, Greece

Available online 4 January 2006

Abstract

Recommender systems are a special class of personalized systems that aim at predicting a user's interest on available products and services by relying on previously rated items or item features. Human factors associated with a user's personality or lifestyle, although potential determinants of user behavior are rarely considered in the personalization process. In this paper, we demonstrate how the concept of lifestyle can be incorporated in the recommendation process to improve the prediction accuracy by efficiently managing the problem of limited data availability. We propose two approaches: one relying on lifestyle alone and another integrating lifestyle within the nearest neighbor approach. Both approaches are empirically tested in the domain of recommendations for personalized television advertisements and are shown to outperform existing nearest neighborhood approaches in most cases.

Keywords: Personalization; Recommenders systems; Collaborative filtering; Content-based filtering; Lifestyle; Digital television; Advertisements

1. Introduction

The vast amount of information available over digital platforms, coupled with the diversity of user information needs, have urged the development of personalized systems

* Corresponding author. Tel.: +30 2108203855; fax: +30 2108203685.
E-mail addresses: glekakos@aueb.gr (G. Lekakos), giaglis@aueb.gr (G.M. Giaglis).
[1] Tel.: +30 2108203682; fax: +30 2108203685.

0953-5438/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.intcom.2005.11.004

that are capable of distinguishing one user from another to provide content, services, and information tailored to individual users. Although personalization has long drawn the attention of academics and practitioners, methods and techniques for implementing personalization remain a topical issue of research. Early typical user modeling approaches, although successful at the academic level, have been proven of little significance in real life applications (Strachan et al., 2000). Conversely, recommender systems, which typically follow low-complexity modeling approaches, have demonstrated remarkable success in actual conditions (Konstan, 2001).

Recommendation approaches typically rely upon implicitly or explicitly acquired behavioral data denoting users' interest but are rarely concerned with factors that characterize the users themselves. Human factors, such as personality and cognitive/learning style, can play an important role in the personalization process and have been studied in user modeling research (Picard, 1997; Hudlicka, 2001; Grabler and Zins, 2002; Fesenmaier et al., 2003). However, lifestyle remains understudied despite being one of the most popular predictors of buying or information seeking behavior in consumer behavior theory.

This paper demonstrates that lifestyle can serve as the basis of effective recommendation methods and proposes novel algorithmic approaches that exploit this factor. The proposed approaches address the sparsity problem, which is considered as one of the most important limitations in several recommendation techniques (as explained below) and occurs when few interaction data are available (Konstan et al., 1997; Balabanovic and Shoham, 1997). Furthermore, the paper captures the parameters associated with the predictive performance of the proposed personalization approaches and explains their superiority.

The application domain of our research is the domain of digital interactive television advertisements, considered as an audiovisual information item that promotes (recommends) products, services, or information to individual viewers (users). The developments in digital interactive television technology (Milenkovic, 1998) provide the ability to replicate (in this environment) methods and techniques applied on other interactive environments and at the same time draw conclusions that can be applicable in other domains as well.

The remaining of the paper is organized as follows. In Section 2, background work is presented, followed by the presentation of the proposed personalization approach in Section 3. The empirical results of the application of the proposed approach are then discussed in Section 4, followed by Section 5 in which different classes of products are examined in terms of the lifestyle approach predictive performance. In Section 6, the proposed approach is extended to demonstrate that its performance can drastically increase within certain hybridization techniques. Conclusions and further research issues are discussed in Section 7.

2. Background

The original idea underlying recommender systems was based on the observation that people very often rely upon opinions and recommendations from friends, family, or associates to make selections or purchase decisions. This 'social' approach to information filtering (Malone et al., 1987) motivated the development of recommender systems defined as systems that 'produce individualized recommendations as output or have the effect of guiding the user in personalized way to interesting or useful objects in a large space of possible options' (Burke, 2002).

2.1. Recommendation methods

The primary concern of recommendation approaches is to identify which of the information items (or objects) available are interesting or likeable to the users. The recommendation process usually takes user ratings on observed items and/or item features as input and produces the same type of output for unobserved items. Ratings can be collected either implicitly, by monitoring the user's interactive behavior (Breese et al., 1998), or explicitly, by asking users to rate the observed items. Two major approaches are used for processing input data and formulating the prediction: collaborative filtering (CF) and content-based filtering (CBF).

Collaborative filtering operates upon the assumption that users who have exhibited similar behavior in the past (or present some form of similarity) can serve as recommenders for each other on unobserved data items. So, given the target user's ratings on observed items, the idea is to trace relationships or similarities between him/her and the remaining of the users in the database, aggregate the 'similar' users' ratings, and use them as a prediction for the target user.

On the other hand, content-based filtering makes predictions upon the assumption that a user's previous preferences or interests are reliable indicators for his/her future behavior. CBF requires that items are described by features, and is typically applied upon text-based documents, or in domains with structured data (Balabanovic and Shoham, 1997; Pazzani, 1999). For example, content-based filtering has been utilized in book recommendation tasks (Mooney and Roy, 2000), using features such as title, author, or theme. In such cases, the user's previous preferences on the respective features are used to filter the available books and recommend the most relevant to the user.

In contrast to content-based filtering, collaborative filtering is applicable to any type of content (Balabanovic and Shoham, 1997), while it can also capture concepts that are hard to represent, such as quality and taste (Herlocker et al., 2002). Additionally, collaborative filtering does not restrict the spectrum of recommendations to items similar to the ones that the user has previously evaluated. Collaborative filtering has been acknowledged as the most successful and most widely implemented recommendation technique to date (Burke, 2002; Sarwar et al., 2000). For these reasons, we will focus on the collaborative filtering strategy in this paper.

The collaborative filtering task concerns the prediction of the target user's rating for a specific item (the target item), based upon his/her ratings on observed items. Each user is represented by item-rating pairs, and can be summarized in a user$\times$item table, which contains the ratings R_{ij} that have been provided by the ith user for the jth item (Table 1).

Collaborative filtering approaches can be distinguished into two major classes: model-based and memory-based (Breese et al., 1998). Model-based methods develop a model, which is applied upon the target user's ratings to make predictions for unobserved items (e.g. Breese et al., 1998; Basu et al., 1998; Billsus and Pazzani, 1998; Pennock et al., 2000; Sarwar et al., 2000). On the other hand, memory-based methods operate upon the entire

Table 1
The user × item representation in collaborative filtering

Users	Items					
	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
User 1	R_{11}	?	R_{13}	R_{14}	R_{34}	R_{16}
User 2	R_{21}	R_{22}	R_{23}	R_{24}	R_{25}	R_{26}
User 3	R_{31}	?	R_{33}	R_{34}	?	R_{36}
User 4	?	R_{42}	?	R_{44}	R_{45}	R_{46}
User 5	R_{51}	R_{52}	R_{53}	?	R_{55}	R_{56}

Question marks indicate missing ratings.

database of users to find the closest neighbors of the target user and weight their recommendation according to their similarities. The fundamental algorithm of the memory-based class is the nearest neighbor (denoted as NN, hereafter), which is considered as one of the most effective collaborative filtering approaches (Good et al., 1999; Herlocker et al., 2002; Schafer et al., 2001; Hofmann, 2004) and can serve as a suitable and reliable algorithm for our recommendation task. It can be described as a process divided in three steps (Resnick et al., 1994):

(a) Measurement of similarities between the target and the remaining users. A typical measure of similarity is the Pearson correlation coefficient (Resnick et al., 1994), which is applied on the items rated in common by two users (Eq. (1))

$$w(i,j) = \frac{\sum_k \left(R_{i,k} - \bar{R}_i\right)\left(R_{j,k} - \bar{R}_j\right)}{\sqrt{\sum_k \left(R_{i,k} - \bar{R}_i\right)^2 \sum_k \left(R_{j,k} - \bar{R}_j\right)^2}}, \tag{1}$$

where $R_{i,k}$ and $R_{j,k}$ refer to the rating of the kth item commonly rated by both users i and j, and $\bar{R}_i$, $\bar{R}_j$ refer to the mean values of the available ratings for the users i and j. The coefficient value ranges from -1 to $+1$, tracing both positive and negative correlations.

(b) Selection of the neighbors who will serve as recommenders. Two techniques have been employed in recommender systems: threshold-based selection (Shardanand and Maes, 1995), according to which users whose similarity exceeds a certain threshold value are considered as neighbors of the target user, and the top-n technique in which a predefined number of n-best neighbors is selected (Resnick et al., 1994).

(c) Prediction based on the weighted average of the neighbors' ratings, weighted by their similarity to the target user:

$$R_{i,p} = \bar{R}_i + \frac{\sum_{j=1}^{m} w(i,j)\left(R_{j,p} - \bar{R}_j\right)}{\sum_{j=1}^{m} |w(i,j)|}, \tag{2}$$

where $R_{i,p}$ is the rating to be predicted for user i and for item p, $\bar{R}_i$ is the mean of the ratings of user i for all items that user has provided his/her ratings, the weight $w(i,j)$ is the similarity measure between user i and j, $R_{j,p}$ is the rating of user j for item p, and $\bar{R}_j$ is the mean of ratings of user j in a neighborhood of size m (the denominator serves as the normalizing factor for the use of the weights in the numerator).

Nearest neighbor algorithms present high-accuracy levels in terms of prediction (Herlocker et al., 2002), confirmed by empirical studies (Breese et al., 1998; Sarwar et al., 2000; Schafer et al., 2001; Basu et al., 1998; Good et al., 1999). The disadvantages of nearest neighbor algorithms are associated with their limited scalability, as the size of the database increases. In order to deal with this problem, heuristics that select only a subset of the original database (Schafer et al., 2001; Hill et al., 1995) or other data mining techniques can be employed (Han and Kamber, 2001).

Nearest neighbor algorithms also present limited explanatory power since recommendations are made upon the overall taste of similar users and not upon specific item features that explain why an item is liked. In addition, nearest neighbor algorithms inherit the limitations of collaborative filtering algorithms in cases of limited availability of user ratings on observed items, which seriously affects the predictive performance, as discussed next.

2.2. The role of lifestyle

The sparsity problem has been acknowledged as the most important drawback in collaborative filtering algorithms (Billsus and Pazzani, 1998; Claypool et al., 1999; Pennock et al., 2000; Chien et al., 1999), and refers to the low ratio of rated items to the total of available items. In general, recommender systems users rate only a small fraction of the available items, since they are not willing to invest time and effort in rating (Aggarwal et al., 1999). Even in systems where ratings are collected implicitly through monitoring of users' interactive behavior, the vast amount of available items and the requirement that users have actually observed and reviewed an item makes the collection of a sufficient number of ratings problematic. In nearest neighbor algorithms, sparsity affects the measurement of similarities, thus leading to unreliable measures and consequently to reduced prediction accuracy (Sarwar et al., 2001). This problem is also significant when a new item is introduced in the system (the 'new item' problem): since no rating exists for the new item, it cannot be recommended to any of the users. Similarly, the 'new user' problem occurs when a new user is introduced in the system, and therefore, no ratings are available for the measurement of his/her similarities with the remaining users.

A major goal in collaborative filtering applications is to improve the accuracy of algorithms by effectively addressing the sparsity problem. We argue that the concept of *lifestyle* can be utilized to overcome this problem and produce more accurate predictions.

Lifestyle is defined as the patterns in which people live and spend their time and money (Gunter and Furnham, 1992). It represents the central notion in the consumer behavior model (Hawkins et al., 1998) which suggests that consumers' actual and desired lifestyle (i.e. the way they would like to think and feel about themselves) are translated into daily behaviors including purchase and consumption behavior. Lifestyle is affected by a number of external (culture, subculture, demographics, social status, reference groups, family, and marketing activities) and internal factors (perception, learning, memory, motives, personality, emotions, and attitudes). It can be quantified through psychographic research that measures constructs revealing attitudes, values and beliefs, interests and activities, demographics, media consumption, and product usage rates. The measurement of these constructs and the application of clustering techniques upon these data lead to the lifestyle segmentation, one of the most effective and popular market segmentation methods (Vyncke, 2002). The clustering process also provides a set of classification rules, which can be applied to consumers' demographic and media consumption data to classify them into the lifestyle segments. Subsequently, the product usage rates attached to the description of the segments are used to infer the preferences of consumers and target products accordingly.

Given the widespread and successful use of lifestyle segments in market targeting activities, we can hypothesize that the concept of lifestyle may also be utilized for the recommendation of products and services, in particular, in sparsity conditions. More specifically, the underlying idea is that, instead of developing the neighborhood of the target user based on unreliable similarities computed upon a few ratings, we can develop a 'lifestyle' neighborhood. Since people can be discriminated upon their lifestyles (Chaney, 1996), and consumers found in the same lifestyle segment (or neighborhood) present common behavior, then the members of a 'lifestyle' neighborhood can be considered as reliable recommenders to each other.

In personalization research, personality traits (such as lifestyle) have been acknowledged as potential personalization factors (Brusilovsky, 2001), but lifestyle has not been adequately studied to date, except in a few cases, such as SeAN (Ardissono et al., 2001) and lifestyle finder (Krulwich, 1997). However, both systems limit the exploitation of lifestyle to the classification of users in lifestyle segments. In research indirectly involving personality or behavioral factors, Pennock et al. (2000) propose a personality diagnosis (PD) algorithm, based on the assumption that there is an association between how people rate items and their personality type (modeled as a latent variable). Other approaches aim at clustering users assuming that behavioral relationships exist among them (Breese et al., 1998; Ungar and Foster, 1998). Besides the fact that such approaches are not concerned with the notion of lifestyle per se, another major difference is that they rely upon available data and are therefore affected by the sparsity problem, while the personality factor upon which clusters are developed is not specified.

In contrast, lifestyle is a meaningful factor that can group users of personalized systems and eventually overcome the sparsity effect. The development of a personalization approach incorporating the notion of lifestyle is presented in Section 3.

3. Developing the lifestyle approach

3.1. The sparsity effect

The main idea underlying the use of lifestyle to overcome the sparsity effect is to avoid considering irrelevant neighbors as relevant (or ignoring actually relevant neighbors) due to the absence of a sufficient number of ratings. Indeed, sparsity affects the most important step in the prediction process, which is the identification of users that will be considered as recommenders (Sarwar et al., 2001). Let us take an example applying the first step of the nearest neighbor algorithm on three users (u_1, u_2, u_3) described by their ratings (on a one-to-five scale) on five items: $u_1 = \{2, 3, 5, 1, 4\}$, $u_2 = \{2, 3, 4, 5, 4\}$, $u_3 = \{2, 1, 1, 3, 2\}$. Computing their similarities (using Eq. (1) of Pearson correlation coefficient) upon five ratings gives quite different results than computing the similarities upon the first three ratings, as depicted in Table 2.

Assume that we are interested in finding the recommenders of u_2, i.e. all users whose similarity with u_2 exceeds a threshold value. Let us take this value to be zero, and therefore, consider all positively correlated users as neighbors. When the users are compared upon three items u_1 is a neighbor while u_3 is not. When more ratings become available u_1 is excluded while u_3 is included in u_2's neighborhood. Who is actually a neighbor is not quite clear. However, we must assume that as the number of items upon which similarities are computed increases similarities become more reliable.

We contend that the incorporation of an additional step in the prediction process, involving the identification of similar users in terms of their lifestyle, will restrict the search space among users that present this form of similarity (lifestyle), thus avoiding the effects of misleading similarity computations. A direct implementation of this additional step is to exploit existing lifestyle segments, relying on their successful utilization in market targeting activities. More specifically, if users are classified into one of these segments then we achieve our objective of restricting the search space within each segment, which includes similar users. Prediction can then be made by assigning the cluster's centroid (the mean rating value of the ratings of the segment members for a specific item) or assigning an expert's predefined prediction (Lekakos and Giaglis, 2004). However, these predictions are assigned uniformly to all users, and therefore, the level of similarity between the target user and his/her neighbor is not taken into account. Thus, an alternative approach would be to weigh each neighbor's contribution in the final prediction by his/her similarity to the target user, following a nearest neighbor reasoning (Lekakos and Giaglis, 2005).

Both options require that users have been pre-classified in lifestyle segments and anchor the whole process to the availability and validity of existing proprietary segments (Mowen and Minor, 1998; Gunter and Furnham, 1992; Beatty et al., 1998). In addition, the classification of users in lifestyle segments is performed under the assumption that each user belongs to one segment only, and therefore behaviorally similar users with different

Table 2
Pearson correlation coefficients computed upon five and three items (positive correlations appear in bold)

Users	u_1	u_2	u_3
Five items			
u_1		0	−0.7559
u_2	0		**0.4193**
u_3	−0.7559	**0.4193**	
Three items			
u_1		**0.9819**	−0.7559
u_2	**0.9819**		−0.8660
u_3	−0.7559	−0.8660	

lifestyles cannot be traced. The development of an approach that may overcome these problems is discussed next.

3.2. The lifestyle approach

To address these limitations we investigate a segment-independent way of identifying lifestyle neighbors directly at the individual level, instead of classifying users into lifestyle segments. This could be achieved by measuring 'lifestyle' similarities directly between the target and the remaining users to develop a 'personal' neighborhood for the target user. Thus, each user must be described by a set of features that meet the following requirements:

(a) They should be lifestyle indicators, i.e. they should be significantly associated with the membership of a user into a lifestyle.
(b) They should be independent from the availability of ratings to reduce the sparsity effect.
(c) They should be easily collectable to avoid increasing the cost of collecting additional information and engaging the users into time-consuming questionnaire filling processes.

Evidence suggests that user demographics can play this role. Indeed, demographics have been successfully used to classify users in *lifestyle finder* (Krulwich, 1997) and *SeAN* (Ardissono et al., 2001). However, demographic data are usually too generic to achieve accurate classification results when used in isolation. For example, in *SeAN*, Ardissono et al. (2001) had to combine demographics with user hobbies, while in cross-selling of banking services (Peltier et al., 2002), demographics were used in conjunction with customer credit data.

Relevant studies in the domain of television advertising have demonstrated that demographics in combination with TV program preferences are significant indicators of membership in lifestyle segments (Lekakos and Giaglis, 2005). In particular, significant discriminating power with respect to lifestyle has demonstrated the combination of the demographic variables 'age', 'marital status', and 'education' with the eight program genres 'documentaries', 'cartoons', 'football/basketball/volleyball games', 'video clips', 'domestic comedy series', 'discussions/interviews', and 'news' (Lekakos, 2004).

These indicators also satisfy the second and third requirements set above. For example, in the domain of digital television, demographics can be collected at the subscription to the service and updated online if necessary, while television program preferences can be easily traced online through the set-top box.

The television program preferences as well as the demographic data are encoded in a uniform binary format and each user is represented by a vector of feature-value pairs. Hence, similarities among users can be computed upon these features and the target user's neighbor consisting of users with 'similar' lifestyle patterns can be traced. While several similarity measures can be applied on binary variables (for example, simple matching, Jaccard coefficient, and phi 4-point correlation) the Pearson correlation coefficient is selected since it has been used in the measurement of similarities upon demographic data

(Pazzani, 1999) as well as for consistency reasons with the nearest neighbor (NN) algorithm that will be used for the evaluation of the proposed approach. Let us call this approach the 'lifestyle' approach, described by the following steps:

1. Measure similarities between the target and the remaining users based upon data associated with their lifestyle by applying Eq. (3)

$$w(i,j) = \frac{\sum_k \left(I_{i,k} - \bar{I}_i\right)\left(I_{j,k} - \bar{I}_j\right)}{\sqrt{\sum_k \left(I_{i,k} - \bar{I}_i\right)^2 \sum_k \left(I_{j,k} - \bar{I}_j\right)^2}}, \tag{3}$$

   where $I_{i,k}$ and $I_{j,k}$, refer to kth lifestyle indicator available in common for the ith (target user) and jth users, and $\bar{I}_i$ and $\bar{I}_j$ to the corresponding means.
2. Formulate the target user's neighborhood, based on the similarity measures described in step 1, by selecting users who score above a certain threshold.
3. Predict the target user's rating on the target item by aggregating lifestyle neighbors' ratings weighted by the lifestyle similarities developed at step 1. Aggregate the target user's preferences into a prediction for the target item by applying Eq. (4) (Resnick et al., 1994; Hill et al., 1995; Shardanand and Maes, 1995):

$$R_{i,p} = \bar{R}_i + \frac{\sum_{j=1}^{m} w(i,j)\left(R_{j,p} - \bar{R}_j\right)}{\sum_{j=1}^{m} |w(i,j)|}, \tag{4}$$

   where—identically to Eq. (2) of the NN approach—$R_{i,p}$ is the rating to be predicted for user i and for item p, $\bar{R}_i$ is the mean of the ratings of user i for all items that user has provided his/her ratings, the weight $w(i,j)$ is the similarity measure between user i and j and $R_{j,p}$ is the rating of user j for item p and $\bar{R}_j$ is the mean of ratings of user j in a neighborhood of size m.

Eq. (3) is the Pearson correlation measure applied on the lifestyle indicators, which are commonly available. In the absence of some of these indicators for a user, the Pearson correlation can be computed upon the remaining indicators. Furthermore, Eq. (4) is identical to the equation used in the NN approach for rating prediction (Eq. (2)). This serves our purpose to take into account the most up-to-date information regarding the preferences of the target user's neighbors as expressed through their ratings. Furthermore, the lifestyle approach (like the NN one) requires that a number of ratings are available for the target item in order to make a prediction.

The main difference between the lifestyle and the NN approach is at the first step of the proposed method where lifestyle indicators, instead of ratings, are used for the computation of similarities. This step results into different neighbors and different weights indicating the importance of each neighbor's rating utilized in the prediction formula. Thus, the prediction accuracy of the proposed approach depends of the validity of the hypothesis that lifestyle neighbors are more reliable than ratings-based neighbors in sparsity conditions. The validity of this hypothesis will be examined in Section 4.

4. Empirical evaluation of the lifestyle approach

4.1. Experimental design

In the experiment, we employed a sample of 37 individuals drawn from our research group. The sample includes academic (19%), research (73%), and technical staff (8%), consisting of 62.2% males and 37.8% females, aged 18–24 (10.8%), 25–34 (67.6%), 35–44 (18.9%) and 45–54 (2.7%). The users were shown 65 advertisements selected from seven product categories (food and drink, fast moving consumer goods, computer and technology, family and home, books and magazines, public services, finance and investment, and autos) and provided their ratings for each advertisement in a one-to-five scale. Similarly to other domains (e.g. movies, books) a rating concerns the degree of likeability (i.e. the overall taste) for each item. Furthermore, participants filled-in a questionnaire providing their demographic and TV program preferences data as required by the lifestyle algorithm.

Empirical findings suggest that the best predictive performance of the NN approach is achieved within a neighborhood of 20–30 users (Herlocker et al., 1999; Herlocker, et al., 2002; Sarwar et al., 2001). Thus, our sample size ensures that the performance of the NN algorithm will not be pessimistic due to a small sample size.

To compare the performance of the two algorithms, a cross-validation technique was used, appropriate for small-sample experiments (Dietterich, 1998). In particular, we employed the leave-one-out (LOO) cross-validation, which is the recommended technique for small-sample model selection problems (Cawley and Talbot, 2003). This method replicates the error estimation process n times for a sample of size n by considering each user in the original sample as the test set (target user) and the remaining sample of size $n-1$ as the training set. A certain number of randomly selected ratings are considered available for each target user to make a prediction for the remaining (removed) ratings. The ratings are removed following the experimental design for the empirical analysis of collaborative filtering algorithms introduced by Breese et al. (1998) who describe a set of experimental protocols, called Given 2, Given 5, and Given 10.The Given n protocol involves the random selection of 2, 5, or 10 votes (corresponding to 'n') from each test user as the observed ratings. These ratings are then used to predict the remaining ratings. The various 'given' protocols examine the performance of the algorithms when relatively little is known about the active user. This set of experimental protocols also includes the 'all-but-one' protocol where a single rating is removed and predicted given all the other ratings. It measures the performance of the algorithms when given as much data as possible (Breese et al., 1998), and therefore, it is beyond the scope of our empirical evaluation, which concerns the performance of the algorithms in sparsity conditions.

In the first step of the NN approach, similarities are computed upon the commonly rated items between the target user and the remaining users. Therefore, all other ratings besides the ones considered available for the target user are actually ignored. In the lifestyle approach, similarities are computed upon the lifestyle indicators. Thus, the number of available ratings does not affect the lifestyle approach in this step, according to our design. Next, a prediction for the target item is made by applying Eqs. (2) and (4) for the NN and lifestyle algorithms respectively, considering users with similarity above a zero threshold value. Taking into account the sample size, the neighbors that actually contribute to the prediction of the target rating may easily drop to a very small number despite our requirement (for both approaches) that the target item has been rated by a sufficient amount of users. Thus, to avoid pessimistic performance of the NN approach, the neighbors' ratings on the target item are also considered available. This design controls the number of ratings utilized in the prediction formula and is applied in both approaches.

The corresponding prediction error for each user is measured using the mean absolute error (MAE), which is the average difference between the predicted and the actual rating value and is commonly used for performance evaluation (Breese et al., 1998; Herlocker et al., 2002; Shardanand and Maes, 1995; Claypool et al., 1999; Melville et al., 2002). The overall error estimator for each of the approaches is the average value of the errors from $(65-n)\times 37$ predictions, where 65 is the total number of items in the dataset, n is the number of available items in each of the 'Given n' experimental protocols, and 37 is the total number of users.

4.2. Experimental results

We focus our attention on the Given 2, 5, and 10 protocols since they represent the sparsity conditions according to our objective. In all cases presented below, the normality requirement is met and the differences in mean absolute errors are compared using paired t-tests (Mitchell, 1997). The ratings of the training and test items follow the same distribution since they are drawn through stratified random sampling according to the distribution of ratings in the intervals A = [1–2), B = [2–3), C = [3–4), D = [4–5] (Basu et al., 1998) to avoid pessimistic performance of the algorithms (Mitchell, 1997).

Further replications were performed (in particular for small values of n such as 2 or 5) to remove possible bias due to the selection of the specific n-ple as the set of training items. The n-ples used for the replications are disjoint (non-overlapping), thus preserving the independence of the test sets (Mitchell, 1997; Dietterich, 1998). In the case of the Given 2 protocol, five replications were performed since 2-ples are randomly selected (stratified selection is not applicable) and as many as 32 disjoint 2-ples are available. For the Given 5 protocol, three replications are sufficient (for the 13 disjoint 5-ples) since training and test items are drawn from the same distribution. The stratified selection of training and test items in the Given 10 protocol, along with the small number of disjoint 10-ples (only 6), allow for a single execution of the experiment.

The averaged results along with the values of the t-statistic are depicted in Table 3.

The results confirm that the *lifestyle* approach achieves lower error levels for all protocols. An interesting finding is that lifestyle approach, although significantly better, is also affected by the number of available items, despite the fact that similarities are computed on the basis of rating-independent data. Considering the prediction formula (Eq. (4)), each user in the target user's neighborhood does not contribute directly by his/her rating but does so by the 'amount' of likeness (positive value of the quantity $R_{j,p}-\bar{R}_j$) or 'dislikeness' (negative value of the above quantity). The weighted sum of the above quantities is added to the mean value of the available ratings of the target user. Thus, if the target user's mean value is misleading with regard to what the user perceives as neutral

Table 3
Overall performance of the algorithms

	Given 2	Given 5	Given 10
Lifestyle	1.1639	0.8265	0.7850
NN	1.1857	0.8416	0.7942
t-value	−5.917	−3.818	−1.946
P	0.0001	0.000	0.05

value due to few ratings, then the prediction may be negatively affected (but still less significantly than in the NN-based prediction).

One of the additional benefits of the lifestyle approach is that it can be further analyzed based on the elements of lifestyle, in contrast to the limited explanatory power of the collaborative filtering approach. In Section 5, we analyze the performance of the lifestyle approach in relation to product-related attributes.

5. Product typologies and prediction accuracy

So far we have treated advertisements and advertised products as the collaborative filtering approach dictates: pieces of information (the term 'item' has been used so far), disregarding their particular characteristics. Computational approaches like collaborative filtering ignore the product itself and cannot provide useful explanations and interpretations concerning possible associations between prediction accuracy and the advertised products. This is a well-known inherent problem of collaborative filtering (Basu et al., 1998; Balabanovic and Shoham, 1997), which restricts the transparency of recommendations, since it offers little knowledge regarding the features that render a specific information item preferable by a user. Furthermore, the superiority of the lifestyle approach when two or five ratings are available has been mainly attributed to the computational disadvantage of the NN approach (at those amounts of given ratings); however, it remains open whether the superiority of the lifestyle approach is associated with specific products. Towards this problem, we explore the Foote, Cone and Belding (FCB) grid (Vaughn, 1980; Ratchford, 1987) that classifies 60 common products along two axes: 'high–low involvement' and 'think–feel'. The idea underlying this section is to utilize these dimensions and compare the performance of the two algorithms at these categorization groups (Table 4).

The FCB grid is a behavior-oriented product typology, which encompasses the typical dimension of high- and low-involvement products that refers to the degree of personal relevance and importance of the product with respect to the purchase consequences or achievement of personal goals (Peter and Olson, 1996). The second dimension lies on the 'think–feel' line, referring to 'think' products with rational meaning, such as functionality (Friedman and Lessig, 1986), and 'feel' products considered by consumers in terms of emotional factors, such as psychosocial consequences and values (Burke and Eddel, 1989).

The following empirical evaluation compares the performances of the lifestyle and NN approaches within the FCB groups to identify possible associations between the performance of the approaches and specific product groups.

Table 4
The FCB grid classifying 60 common products into four classes

	Think	Feel
High	Informative	Affective
	Life/auto insurance, family car, washer, camera etc.	Sport car, expensive watch, perfume etc.
Low	Habit-formation	Self-satisfaction
	Bleach, suntan lotion, razors, salad oil etc.	Pizza, beer, soft-drinks, magazines etc.

Table 5
Distribution of products in the FCB groups in a total of 26 products

	High	Low	Total (think, feel)
Think	8 (13)	7 (23)	15
Feel	4 (4)	7 (25)	11
Total (high, low)	12	14	26

Figures in parentheses refer to the original distribution of products in the 65-item dataset.

The first step for the evaluation task is preprocessing the original 65-item data set to achieve a more uniform distribution of products among the FCB groups and ensure the balanced representation of all products in the experiments. This pre-processing step results into a 26-item dataset that meets this requirement (Table 5).

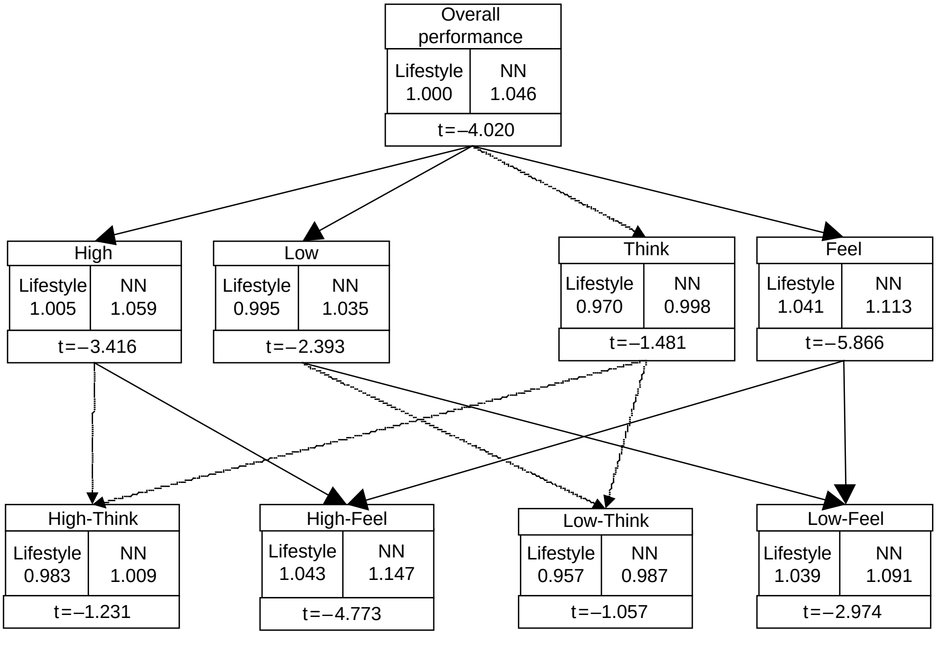

Fig. 1. The associations between main and secondary FCB categorization groups for the Given 2 protocol.

In terms of experimental design, we follow the Breese et al. (1998) methodology (same observed, same test items) to control the stratified selection of products for the training and test sets. We first measure the performance along the main axes 'high–low' and 'think–feel' for each protocol and then we proceed in the analysis of those differences within the four combined groups ('high–think', 'high–feel', 'low–think', and 'low–feel').

Fig. 1 depicts the paths along which significant differences can be found for each combination in the Given 2 protocol. Dashed arrows connect main groups ('high', 'low', 'think', 'feel') to secondary groups (combinations of the main groups) with non-significant differences, while bold arrows indicate a path where groups with significant differences are found. The values of the *t*-statistics are also presented along with the MAEs for each case.

As the above results suggest the superiority of the lifestyle approach (Given 2 protocol) in the averaged performance is also reflected in its superiority on both the high-involvement and low-involvement products. In addition, the associations between the FCB categorization groups for the Given 2 protocol indicate the role of the 'feel' factor on the superiority of the lifestyle approach. Indeed, the combination of 'high–feel' and 'low–feel' features presents significant differences in favor of the lifestyle approach and can explain the differences within the 'high', 'low', and 'feel' groups. It is also clear from the value of the *t*-statistic (−4.773) that the most powerful combination for the *lifestyle* approach is the 'high–feel' group, which includes products such as sport cars that represent an important decision for the consumer and at the same item raise emotional decision factors.

Following the same analytical approach for the Given 5 protocol (Fig. 2), we can infer that lifestyle outperforms the NN approach for high-involvement products.

It is interesting that, despite the fact that no significant differences can be found for the 'feel' products, the combination 'high–feel' confirms the findings of the Given 2 protocol

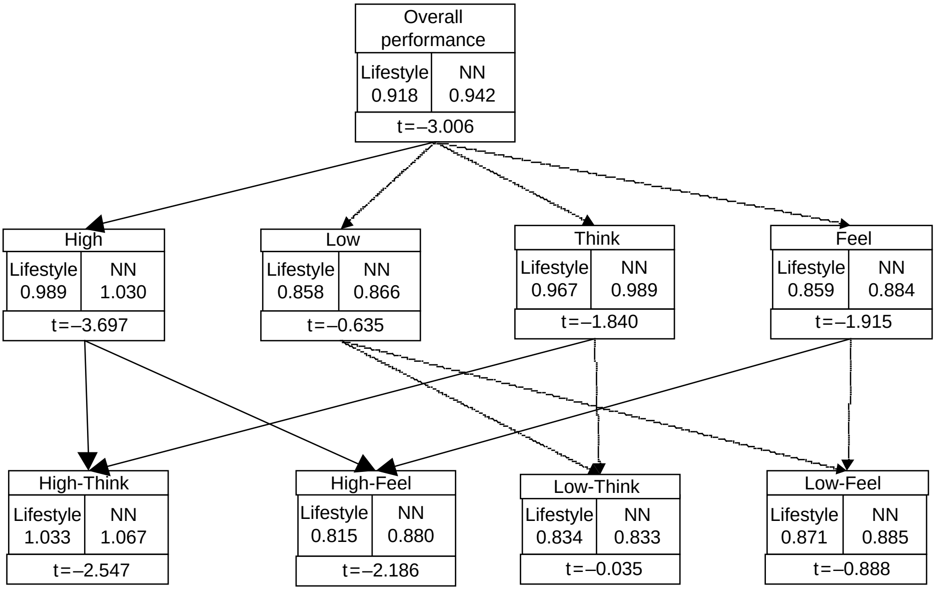

Fig. 2. The associations between main and secondary FCB categorization groups for the Given 5 protocol.

and gives significant differences in favor of the lifestyle approach. In addition, for the Given 5 protocol the 'high–think' combination also gives significant differences, and therefore, the superiority of the lifestyle approach for the high-involvement products can be attributed partially to the 'think' and partially to the 'feel' factor.

These findings can provide, to a certain extent, explanations of the lifestyle approach superiority on the Given 2 and 5 protocols. It is confirmed on both protocols that the lifestyle neighbors utilized in this approach are reliable recommenders, in particular for 'feel' products. However, analogous differences in the performances of the two approaches cannot be found as more items become available since the NN approach increases its performance.

The FCB grouping provides a suitable way to manage the study of a wide range of products. However, the examination of the role of products and product categories upon the future preferences of an individual user lies beyond the boundaries of the collaborative filtering framework and requires the extension of the current study to a content-based analysis.

6. Extending the lifestyle approach

So far we have considered, measured, and compared the performance of individual algorithms and concluded that the lifestyle approach outperforms the NN one in the absence of a sufficient amount of ratings for each user. Exploiting the behavior of the two algorithms, in this section, we apply an 'integrated' recommendation mechanism that utilizes the lifestyle approach to make predictions upon the (eventually few) available ratings for the items unobserved by the target user, populating his/her ratings vector. A new vector is introduced for each user, consisting of the original ratings provided by the user and the ratings predicted by the application of the lifestyle approach. Thus, for each user a new user representation is developed, called 'pseudo-user', who participates as a significant neighbor in a subsequent application of the NN prediction process, which then produces the final prediction for the unobserved items (Fig. 3).

More specifically, let us assume a number u of users where each user_i, $i=1,\ldots, u$ has provided his/her ratings for k_i items $\{R_{i,1}, R_{i,2}, \ldots R_{i,k_1}\}$ in a total of n items, and prediction concerns a rating $R_{t,p}$ for a user_t (target user) and an item p (target item). Then the steps of the approach can be described as follows:

Lifestyle prediction

1. For each pair of users (user_i, user_j), $\text{i,j}=1,\ldots, \text{u}$, $\text{i}\neq\text{j}$ measure their similarity w(i,j) upon their lifestyle indicators, applying Eq. (3) (Section 3.2).
2. For each user_i, $\text{i}=1,\ldots, \text{u}$, develop his/her neighborhood (NB_i) as follows: $\text{NB}_\text{i}=\{\text{user}_\text{j} \mid \text{w(i,j)}>\text{threshold}\}$.
3. For each user_i predict the $\text{n}-\text{k}_\text{i}$ missing ratings $\{L_{i,1}, L_{i,2}, \ldots, L_{i,n-k_i}\}$ utilizing (in Eq. (4)) the available ratings from all $\text{users}_\text{j}\in\text{NB}_\text{i}$ for each of the above missing ratings.

Pseudo-user table formulation

4. For each user_i, introduce in the pseudo-user table a pseudo-user p_user_i, whose ratings $\text{P}_\text{i,q}$ are defined as follows:

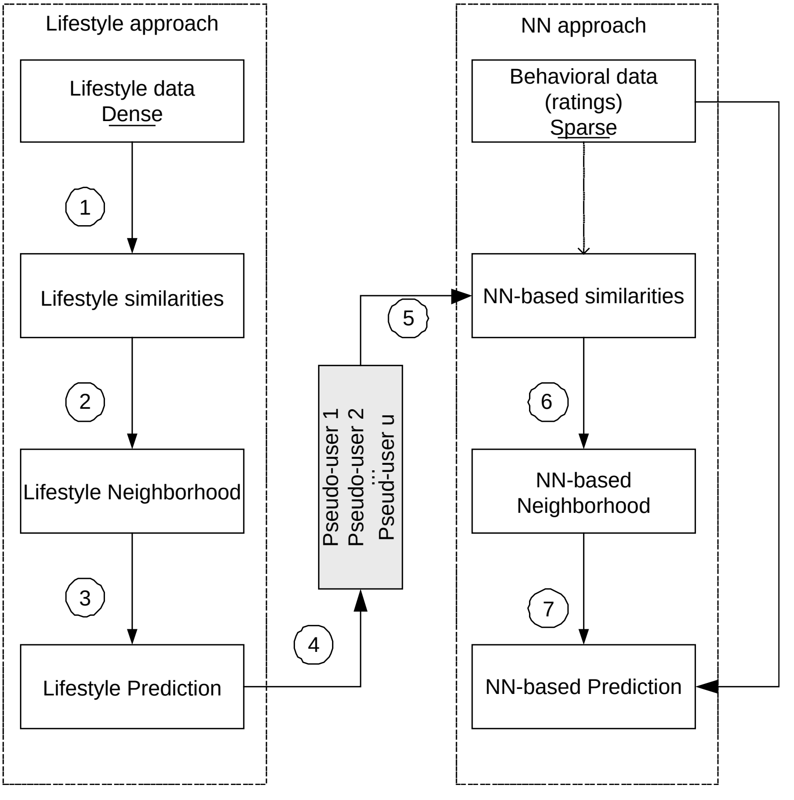

Fig. 3. The 'integrated' personalization approach.

$$P_{i,q} = \begin{cases} R_{i,q}, & \text{if rating for item } q \text{ has been provided by user } i \\ L_{i,q}, & \text{if item } q \text{ has not been rated by user } i \end{cases}, \quad i = 1,\ldots,u, \quad q = 1,\ldots,n$$

and $L_{i,q}$ refers to the predicted rating by the lifestyle approach.

NN-based prediction

5. Measure similarities between the target user_t and each of the remaining users_j, $j = 1,\ldots, u$, $j \neq t$, using the Pearson correlation formula (Eq. (1)). Assign an increased weight to the p_user_t (whose k_t ratings are identical with the target user's ratings, and therefore, we may be more confident for his/her recommendations on the missing ratings).
6. Develop the target user's neighborhood by selecting users above a threshold value.
7. Produce the prediction by the NN prediction formula (Eq. (2)) for the neighbors selected in step (6) and weights computed in step (5).

It is clear that, following the above process, the original sparse data table is removed and all computations are performed upon a 100% dense pseudo-user $\times$ item table as indicated in Fig. 4 (mark (●) denotes originally available rating, ($\times$) a missing rating and (◘) a predicted rating by the individual lifestyle approach).

Original Table						
User 1	•	×	•	×	×	×
User 2	•	×	×	•	×	×
User 3	×	×	×	×	•	•
User 4	×	•	×	×	•	×
User k-1	•	×	×	•	×	×
Target	•	×	•	×	×	×

↓

Pseudo-user Table						
Pseudo-User 1	•	□	•	□	□	□
Pseudo-User 2	•	□	□	•	□	□
Pseudo-User 3	□	□	□	□	•	•
Pseudo-User 4	□	•	□	□	•	□
Pseudo-User k-1	•	□	□	•	□	□
Target	•	×	•	×	×	×
Pseudo-Target	•	□	•	□	□	□

Fig. 4. The transfer of the original sparse data table into a dense pseudo-user table.

The utilization of the lifestyle prediction is very promising, since sparsity is completely removed but the lifestyle predictions also transfer the errors produced by the lifestyle approach and the effect of massive substitution of all users by their respective pseudo-users remains to be investigated. The main hypothesis to be tested is that the integrated approach significantly outperforms the NN approach. The evaluation results will also demonstrate the differences in terms of performance between the integrated approach and the lifestyle and NN approaches discussed so far.

6.1. Empirical evaluation

The design utilizes once again the ‘Given *n*’ protocols, upon the sparse user × item table. For each user the same number of randomly selected ratings is removed and both algorithms are applied upon the ‘original table’ depicted in Fig. 4 (for the Given 2 protocol). Note that only originally available items are considered by both approaches and utilized in the prediction of the missing ratings and the formation of the pseudo-user table. At this phase of the experimentation series, we extend the number of given ratings beyond the Given 2, 5, and 10 ratings. Specifically, the performance of all algorithms discussed so far is measured upon 2, 5, 10, 15, 20, 25, 30, 35, and 50 ratings, to acquire an overall picture of the performance of the various approaches in different levels of available ratings. In the approaches discussed so far, the focus was on proving their superiority to the NN approach upon few ratings. In addition, the experimental results demonstrated that the NN approach improves its performance as more ratings become available. However, our expectations for improved performance of the integrated approach at any sparsity level (compared to the NN), suggest the extension of the experiment in various amounts of available ratings.

Table 6
Predictive performance of the approaches on different sparsity levels

	NN	Lifestyle	Integrated
G2	**1.1071**	1.1053	1.1048
P	0.010	0.142	
G5	**1.0056**	**0.9881**	0.9783
P	<0.0001	<0.0001	
G10	**0.9333**	**0.9286**	0.9057
P	<0.0001	<0.0001	
G15	**0.906**	**0.9059**	0.8776
P	<0.0001	<0.0001	
G20	**0.8859**	**0.8884**	0.8556
P	<0.0001	<0.0001	
G25	**0.8679**	**0.8653**	0.8309
P	<0.0001	<0.0001	
G35	**0.8352**	**0.8381**	0.8081
P	<0.0001	<0.0001	
G50	**0.7784**	**0.7867**	0.7687
P	0.045	0.006	

MAE's in bold indicate significant difference with respect to the integrated approach.

Two replications on different training sets, randomly selected from each user's available ratings have been performed for the cases of 2, 5, 10, and 15 ratings. For the rest of the cases (20 through 35) a single run of the experiment was performed since any subsequent randomized selection of training 20-ples, 30-ples and 35-ples of ratings would result in highly overlapping training sets with no contribution in the final result. Finally, for the case of 50 available ratings, three replications of the experiment are performed on three different (mutually exclusive) sets of test items.

The hypothesis of significant difference of the performance between the integrated and the NN approach is tested by paired *t*-tests between the performances of the two algorithms for each individual user. In addition to the main hypothesis, the integrated approach is also compared to the lifestyle approach to assess the value and usefulness of the integrated approach (Table 6).

The results clearly suggest that the integrated approach is significantly better than all the approaches examined so far, apart from the case of Given 2, where the availability of only two ratings has significant and rather unmanaged effects on all approaches. In the case of the 'Given 50' protocol, the significance levels of the integrated approach superiority are downscaled, compared to the rest of the experiments—as expected—but still the averaged performance is significantly better.

6.2. Discussion

A first observation is that all approaches increase their performance as more items are added in the training test, confirming the theoretical expectation concerning the sparsity effect upon the performance of a learning algorithm.

In contrast to the lifestyle approach, the integrated approach firmly outperforms the NN in the range of 5–35 available ratings, while the improvement is decreased for 50 ratings. At

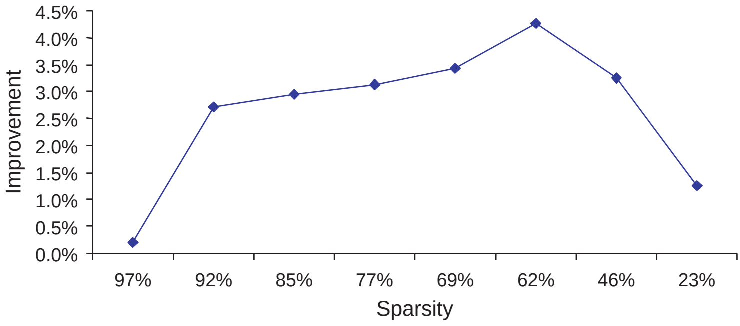

Fig. 5. Relative improvement of the integrated approach.

the Given 2 protocol, all approaches are affected by the limited number of available ratings and present similar performances (although statistical significance is confirmed for the difference between the integrated and the NN approach). Measuring the improvement in performance by the difference in MAE values between the NN and integrated approaches in relation to the various sparsity levels (which correspond to the number of available ratings), it is found that the improvement rate escalates from a very small 0.2% at 97% sparsity (Given 2) to a 2.7% improvement rate (Given 5). Then, it follows an incremental route up to a 4.3% at 62% sparsity level (Given 25), which is the maximum rate. The improvement still maintains a high rate of 3.2% for the Given 35 protocol (46% sparsity) and finally drops at 1.2% when sparsity is only at 23% and the NN approach presents its best performance (Fig. 5). This trend indicates that as we approximate a completely dense user $\times$ item table the NN approach would perform better than the integrated approach.

The performance of the integrated approach compared to the NN one suggests that the management of the sparsity effect can significantly improve the accuracy of the prediction, within the limits set out by the errors inherited by the two approaches upon which the integrated approach has been built. However, the value of the integrated approach lies in its flexibility to accommodate any improved variation of the lifestyle and/or NN approaches. Furthermore, a second important feature of the approach is that it is open to any type of prediction mechanism that is affected by sparse datasets. Indeed, the utilization of lifestyle data and the prediction of missing ratings remove the sparsity effect. This type of outcome (predicted ratings) can be subsequently loaded as input to virtually any learning algorithm (for example, Bayesian Networks) and improve its performance.

7. Conclusions and future research

User personality factors, such as lifestyle, are rarely considered in the personalization process despite their theoretical significance in the prediction of a user's future behavior. We have proposed personalization approaches based on the notion of lifestyle, which has been shown to increase the accuracy of prediction in sparsity conditions, compared to established recommendation approaches, such the nearest neighbor algorithm. The main idea underlying the development of the lifestyle approach is to avoid erroneous selection of neighbors due to misleading computation of similarities among users in sparse databases. The lifestyle approach can overcome this problem by performing the prediction task within the range of lifestyle neighbors. In addition to the computational advantage, the superiority of the lifestyle approach in high-sparsity environments can be attributed to its ability to better predict the behavior of users against 'feel' or 'high-involvement' products (in the domain of television advertisements, where the approaches have been empirically tested).

Moreover, we have followed a hybridization technique based on the notion of pseudo-user, to strengthen the performance of the lifestyle approach and we have demonstrated that the integrated approach yields superior predictive performance at varying sparsity levels. This is achieved by combining the ability of the lifestyle approach to make predictions when few ratings are available with the ability of the nearest neighbor approach to make successful predictions upon sufficient amounts of available ratings.

However, a pre-selection of a sub-sample or the application of some other data reduction algorithm is required to apply the integrated approach in large databases due to the increased computational cost of the method. Furthermore, our empirical findings may be limited by the sample size and the population from which the sample was drawn, which consists mainly of well-educated IT literate individuals. Sample size limitations have been managed to a certain extent by applying cross-validation techniques and exploiting particular features of nearest neighbor algorithms concerning the optimal neighborhood size. However, our future research plans include the evaluation of the proposed approaches upon larger datasets from different domains.

One of the challenges emanating from the present research is to apply the proposed personalization approaches in other domains, such as personalized systems over mobile platforms, which are also characterized by limitations in collecting extended user-driven interaction data. The identification of lifestyle indicators, which are easily collectable in other environments, or even the complete disengagement of the process from such indicators would also increase the generalization ability of the lifestyle approach. In addition, the combination of the lifestyle approach as a collaborative filtering method with content-based filtering could further increase the accuracy and transparency of predictions. Although the proposed approaches address the sparsity problem they cannot operate and make predictions for new items (that have not been rated) and therefore the exploitation of existing relationships between items would eventually lead to a personalization approach that operates in any situation related to the absence of ratings.

References

Aggarwal, C., Wolf, J., Kun-Lung, W., Yu, P.S., 1999. Horting hatches an egg: a new graph-theoretic approach to collaborative filtering. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, San Diego, CA, pp. 201–212.

Ardissono, L., Console, L., Torre, I., 2001. An adaptive system for the personalized access to news. AI Communications 14 (3), 129–147.

Balabanovic, M., Shoham, Y., 1997. Fab: content-based collaborative recommendation. Communications of the ACM 40 (3), 66–72.

Basu, C., Hirsh, H., Cohen, W., 1998. Recommendation as classification: using social and content-based information in recommendation. In: Proceedings of the 15th National Conference on Artificial Intelligence, Madison, WI, pp. 714–720.

Beatty, S.F., Homer, P.M., Kahle, L.R., 1998. Problems with vals in international marketing research: an example from an application of the empirical mirror technique. Advances in Consumer Research 15, 375–380.

Billsus, D., Pazzani, M., 1998. Learning collaborative information filters. In: Proceedings of the 15th National Conference on Machine Learning, pp. 46–54.

Breese, J.S., Heckerman, D., Kadie, C., 1998. Empirical analysis of predictive algorithms for collaborative filtering. In: Proceedings of the 14th Annual Conference on Uncertainty in Artifical Intelligence, pp. 43–52.

Brusilovsky, P., 2001. Methods and techniques of adaptive hypermedia. User Modeling and User Adapted Interaction 6 (2–3), 87–129.

Burke, R., 2002. Hybrid recommender systems: survey and experiments. User Modeling and User Adapted Interaction 12, 331–370.

Burke, M.C., Eddel, J.A., 1989. The impact of ad-based affect and cognition. Journal of Marketing Research 26 (2), 69–83.

Cawley, G.C., Talbot, N.L.C., 2003. Efficient leave-one-out cross-validation of kernel fisher discriminant classifiers. Pattern Recognition 36 (1), 2585–2592.

Chaney, D., 1996. Lifestyles. Routledge, London.

Chien, Y.-H., George, E., 1999. A bayesian model for collaborative filtering. In: Proceedings of the Seventh International Workshop on Artificial Intelligence and Statistics, Fort Lauderdale, Florida, http://uncertainty99.microsoft.com/chien.htm.

Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., Sartin, M., 1999. Combining content-based and collaborative filters in an online newspaper. In: Proceedings of the ACM SIGIR Workshop on Recommender Systems, Berkeley, CA, http://www.csee.umbc.edu/~ian/sigir99-rec/.

Dietterich, T.G., 1998. Approximate statistical tests for comparing supervised classification learning algorithms. Neural Computation 10, 1895–1923.

Fesenmaier, D.R., Ricci, F., Schaumlechner, E., Wober, K., Zanella, C., 2003. Dietorecs: travel advisory for multiple decision styles. In: Proceedings of the Information and Communication Technologies in Tourism, pp. 232–241.

Friedman, R., Lessig, V.P., 1986. A framework of psychological meaning of products. Advances in Consumer Research 13, 338–342.

Good, N., Schafer, J.B., Konstan, J., Borchers, A., Sarwar, B., Herlocker, J., Riedl, J., 1999. Combining collaborative filtering with personal agents for better recommendations. In: Proceedings of the 16th National Conference on Artificial Intelligence (AAAI-99), Menlo Park, CA, pp. 439–446.

Grabler, K., Zins, A., 2002. Vacation trip decision styles as basis for an automated recommendation system: lessons from observational studies. In: Proceedings of the ENTER Conference, Innsbruck, Austria.

Gunter, B., Furnham, A., 1992. Consumer Profiles: An Introduction to Psychographics. Routledge, London.

Han, J., Kamber, M., 2001. Data Mining Concepts and Techniques. Morgan Kaufmann, San Francisco, CA.

Hawkins, I., Best, R.J., Coney, K.A., 1998. Consumer Behavior: Building Marketing Strategy. Irwin, New York.

Herlocker, J., Konstan, J., Borchers, A., Riedl, J., 1999. An algorithmic framework for performing collaborative filtering. In: Proceedings of the 22nd International Conference on Research and Development in Information Retrieval (SIGIR '99)), New York, pp. 230–237.

Herlocker, J., Konstan, J., Riedl, J., 2002. An empirical analysis of design choices in neighborhood-base collaborative filtering algorithms. Information Retrieval 5, 287–310.

Hill, W., Stead, L., Rosenstein, M., Furnas, G., 1995. Recommending and evaluating choices in a virtual community of use. In: Proceedings of the ACM CHI '95 Conference on Human Factors in Computing Systems, pp. 194–201.

Hofmann, T., 2004. Latent semantic models for collaborative filtering. ACM Transactions on Information Systems 22 (1), 89–115.

Hudlicka, E., 2001. Integrated model of trait and state: influences on user behavior. In: Proceedings of the Second Workshop on Attitude, Personality and Emotions in User-Adapted Interaction, in conjunction with User Modeling 2001, Sonthofen, Germany.

Konstan, J., 2001. Heavyweight applications of lightweight user models: a look at collaborative filtering, recommender systems and real-time personalization. In: Proceedings of the Eigth International Conference in User Modeling, UM 2001, Sonthofen, Germany, p. 314.

Konstan, J., Miller, B., Maltz, D., Herlocker, J., Gordon, L.R., Riedl, J., 1997. Applying collaborative filtering to usenet news. Communications of the ACM 40 (3), 77–87.

Krulwich, B., 1997. Lifestyle finder: intelligent user profiling using large-scale demographic data. AI Magazine, 37–45.

Lekakos, G., 2004. Personalized advertising services through hybrid recommendation methods: the case of digital interactive television. PhD Thesis, Athens Univesity of Economics & Business, Athens, Greece.

Lekakos, G., Giaglis, G.M., 2004. A lifestyle-based approach for delivering personalized advertisements in digital interactive television. Journal of Computer Mediated Communication 9 (2).

Lekakos, G., Giaglis, G., 2005. Personalization of advertisements in the digital TV context. In: Shen, C., Magoulas, G. (Eds.), Adaptable and Adaptive Hypermedia Systems. Idea Group, pp. 264–283.

Malone, T.W., Grant, K.R., Turbak, F.A., Brobst, S.A., 1987. Intelligent information systems. Communications of the ACM 30 (5), 390–402.

Melville, P., Mooney, R.J., Nagarajan, R., 2002. Content-boosted collaborative filtering for improved recommendations. In: Proceedings of the 18th National Conference on Artificial Intelligence (AAAI), Edmonton, Canada, pp. 187–192.

Milenkovic, M., 1998, Delivering Interactive Services via a Digital TV Infrasturcture. IEEE Multimedia, 5 (4), 34–43.

Mitchell, T., 1997. Machine Learning. McGraw-Hill.

Mooney, R.J., Roy, L., 2000. Content-based book recommending using learning for text categorization. In: Proceedings of the Fifth ACM Conference in Digital Libraries, San Antonio, TX, pp. 195–204.

Mowen, J.C., Minor, M., 1998. Consumer Behavior, fifth ed. Prentice-Hall, Upper Saddle River, NJ.

Pazzani, M., 1999. A framework for collaborative, content-based and demographic filtering. Artificial Intelligence Review 13 (5–6), 393–408.

Peltier, J.W., Schibrowsky, J.A., Schultz, D.E., Davis, J., 2002. Interactive psychographics: cross-selling in the banking industry. Journal of Advertising Research 42 (2), 7–22.

Pennock, D.M., Horvitz, E., Lawrence, S., Giles, L., 2000. Collaborative filtering by personality diagnosis: a hybrid memory and model-based approach. In: Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence, San Francisco, pp. 473–480.

Peter, P.J., Olson, J.C., 1996. Consumer Behavior and Marketing Strategy, fourth ed. Irwin, Homewood, IL.

Picard, R.W., 1997. Affective Computing. MIT Press, Cambridge, MA.

Ratchford, B.T., 1987. New insights about the fcb grid. Journal of Advertising Research 27 (4), 24–38.

Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., Riedl, J., 1994. Grouplens: an open architecture for collaborative filtering of netnews. In: Proceedings of the ACM Conference on Computer Supported Cooperative Work, pp. 175–186.

Sarwar, B., Karypis, G., Konstan, J., Riedl, J., 2000. Analysis of recommendation algorithms for e-commerce. In: Proceedings of the ACM E-commerce Conference, pp. 158–167.

Sarwar, B., Karypis, G., Konstan, J.A., Riedl, T., 2001. Item based collaborative filtering recommendation algorithms. In: Proceedings of the Tenth International World Wide Web Conference, pp. 285–295.

Schafer, J.B., Konstan, J.A., Riedl, J., 2001. Electronic-commerce recommender systems. Journal of Data Mining and Knowledge Discovery 5 (1), 115–152.

Shardanand, U., Maes, P., 1995. Social information filtering: algorithms for automating 'word of mouth'. In: Proceedings of the ACM CHI'95 Conference on Human Factors in Computing Systems, Denver, Colorado, pp. 210–217.

Strachan, L., Anderson, J., Murray, S., Evans, M., 2000. Minimalist user modelling in a complex commercial software system. User Modeling and User Adapted Interaction 10, 109–145.

Ungar L., Foster, D., 1998. Clustering methods for collaborative filtering. In: Proceedings of the Workshop on Recommendation Systems at the Fifteenth National Conference on Artifical Intelligence, Madison, Wisconsin, pp. 112–128.

Vaughn, R., 1980. How advertising works: a planning model. Journal of Advertising Research 20 (5), 27–33.

Vyncke, P., 2002. Lifestyle segmentation: from attitudes, interests and opinions, to values, aesthetic styles, life visions and media preferences. European Journal of Communication 17 (4), 445–464.