

User Modeling and Recommendation Techniques for Personalized Electronic Program Guides

Liliana Ardissono, Cristina Gena, Pietro Torasso

Dipartimento di Informatica, Università di Torino, Corso Svizzera 185, 10149 Torino, Italy

(email: {liliana, cgena, torasso}@di.unito.it)

Fabio Bellifemine, Angelo Difino, Barbara Negro

Telecom Italia Lab, Multimedia Division, Via G. Reiss Romoli 274, 10148 Torino, Italy

(email: {bellifemine, difino, negro}@tilab.com)

Abstract. This chapter presents the recommendation techniques applied in Personal Program Guide (PPG). This is a system generating personalized Electronic Program Guides for Digital TV. The PPG manages a user model that stores the estimates of the individual user's preferences for TV program categories. This model results from the integration of different preference acquisition modules that handle explicit user preferences, stereotypical information about TV viewers, and information about the user's viewing behavior. The observation of the individual viewing behavior is particularly easy because the PPG runs on the set-top box and is deeply integrated with the TV playing and the video recording services offered by that type of device.

1. Introduction

With the expansion of TV content, digital networks and broadband, hundreds of TV programs are broadcast at any time of day. This huge amount of content has the potential to optimally satisfy individual interests, but it makes the selection of the programs to watch a very lengthy task. Therefore, TV viewers end up watching a limited number of channels and ignoring the other ones; see Smyth and Cotter (in this volume) for a discussion about this issue.

In order to face the information overload and facilitate the selection of the most interesting programs to watch, personalized TV guides are needed that take individual interests and preferences into account. As recommender systems have been successfully applied to customize the suggestion of items in various application domains, such as e-commerce, tourism and digital libraries (Resnick and Varian, 1997; Riecken, 2000; Mostafa, 2002), several efforts have been recently made to apply this technology to the Digital TV world. For instance, collaborative filtering has been applied in the MovieLens (2002) and in the PTV Listings Service (Cotter and Smyth, 2000) systems to generate personalized TV listings, and in the TiVo (2002)

system to select programs for VCR recording. Collaborative filtering requires that the user positively or negatively rate the programs she has watched; the ranking profiles are collected in a central server and clustered to identify people having similar tastes. When somebody asks for a recommendation, the system suggests those items that have been positively rated by the users with the most similar profiles.

Although collaborative filtering suits Web-based applications in an excellent way, we believe that personalized EPGs should rely on recommendation techniques that can be applied locally to the user's TV. In fact, an EPG embedded in the set-top box may continuously track the user's viewing behavior, unobtrusively acquiring precise information about her preferences. Moreover, the guide can be extended to become a personal assistant helping the user to browse and manage her own digital archive.

To prove our ideas, we developed the Personal Program Guide (PPG). This is a personalized EPG that customizes the TV program recommendation and assists the user in the retrieval of the programs she has recorded. The PPG runs on the user's set-top box and downloads information about the available TV programs from the satellite stream. In order to obtain precise estimates of the individual TV viewer's preferences during the whole lifecycle of the EPG, our system relies on the management of a hybrid user model that integrates three sources of information:

- The user's explicit preferences that may be declared by the user.
- Information about the viewing preferences of stereotypical TV viewer classes.
- The user's viewing behavior.

The system customizes the recommendation of TV programs by taking the user's preferences for TV program categories and channels into account. The combination of these two types of information supports accurate suggestions. In fact, the program categories preferred by the user may be privileged. For instance, movies might be recommended more frequently than documentaries. Moreover, within each category, the individual programs selected by the content providers may be prioritized on the basis of their audience analysis.

While the multi-agent architecture of the PPG has been described in (Ardissono et al., 2003), this chapter presents the recommendation techniques applied in the system. The chapter also presents the results of a preliminary evaluation of the PPG with real users. More specifically, Section 2 outlines the facilities offered by the PPG and sketches the representation of the information about TV programs. Section 3 presents the management of the user models. Section 4 describes the recommendation techniques applied to personalize the suggestion of TV programs. Section 5 reports the results of the system evaluation and Section 6 compares our approach to the related work. Finally, Section 7 concludes the paper and outlines our future work.

2. Overview of the Personal Program Guide

The PPG offers advanced facilities for browsing TV content. For instance, the user can search programs by channel, category, viewing time, day, language and cast; see the buttons located in the left portion of the User Interface shown in Figure 1.



Figure 1. User Interface of the Personal Program Guide (PC Simulator)

Moreover, the user may ask for details about a program (e.g., cast, content description and parental rating), she can record it, ask to be advised when the transmission of the program starts (memo function), and so forth. The user can also retrieve the list of programs she has asked to be alerted about (Memo TV events), she has recorded (Recorded TV Events button), or she has bought (Bought TV Events). Although the system acquires the information about the user's interests in an unobtrusive way, it also accepts explicit feedback about programs that may be rated by clicking on the "thumb up/down" buttons located in the bottom-right area of the User Interface.

By default, the system works in personalized mode (Personalization ON) and ranks the TV programs by taking the user model into account. The less suitable programs

are filtered out and the most promising ones are shown at the top of the list. The recommendation degree of a program is represented by a list of smiling faces close to its description in order to make the ranking information independent of the visualization criterion. The personalization facility can be switched off and in that case the TV programs are sorted on the basis of their starting time.

As described in Ardissono et al. (2001), the information about TV programs is based on an extension of the Digital Video Broadcasting standard (DVB, 2000). A record whose fields specify information such as the starting time, the transmission channel and the stream content, i.e., video, audio or data, describes each TV program. The descriptor includes one or more program categories (*Content* field) representing the program content and format. The program categories are organized in the General Ontology, a taxonomy that includes broad categories, such as Serial, and specializes them in sub-categories, e.g., Soap Opera and Science Fiction Serial.

3. A Hybrid User Model for the Specification of TV Viewing Preferences

In the design of the user model, we considered:

- Explicit *preferences* for TV program categories that the user notifies the system about; e.g., movies and documentaries.
- *Estimates on the viewing preferences* for the program categories. These are related to the number of programs she watches, for each category.
- *Socio-demographic information*, such as her age, occupation, and so forth.
- Information about the user's general *interests, hobbies and lifestyles*.
- *Prior information about the preferences of stereotypical classes of TV viewers*.

In order to manage suitably this heterogeneous information, we designed the User Modeling Component (UMC) of the PPG as an agent that exploits three modules, the Explicit Preferences Expert, the Stereotypical UM Expert and the Dynamic UM Expert, each one managing a private user model.

- The *Explicit User Model* stores the information elicited from the user.
- The *Stereotypical User Model* stores the prediction on the user's preferences inferred from prior information about TV viewer categories.
- The *Dynamic User Model* stores the estimates on the user's preferences inferred by observing her viewing behavior.

The predictions generated by the Experts may be affected by uncertainty, e.g., because they have been made in the presence of limited information about the user. In order to take this fact into account, the *confidence* of each prediction is evaluated. The UMC employs this parameter to weight the predictions provided by the Experts into a *Main User Model*, whose contents are exploited to personalize the suggestion of TV programs.

3.1. THE EXPLICIT USER MODEL

This user model stores the user's personal data, (e.g., occupation and age), her declared attitudes towards topics such as cinema, books and politics (henceforth, *general interests*), and her preferences for TV program categories. The system acquires this information by means of a form filled in at registration time.¹ The user may express her interests and preferences by choosing between three values (low, medium, strong) that correspond to numerical values in the user model (0, 0.5, 1).

In order to limit the overhead on the user, the information about her preferences is elicited on few, broad program categories. As these categories are less detailed than those of the General Ontology, suitable mappings between the concepts are defined to enable the inference of the user's preferences.

A confidence value is associated to each prediction to represent the possible uncertainty of the information. The confidence is a decimal number in [0,1], where 0 represents the total lack of confidence and is associated to unknown preferences. The 1 value denotes maximum confidence and is associated to the preferences for the categories of the General Ontology that coincide with the declared user preferences.

3.2. THE STEREOTYPICAL USER MODEL

3.2.1. Representation of the Stereotypical Information

A knowledge base stores the information about TV viewer classes that are represented as stereotypes (Rich, 1989). We defined the stereotypes by exploiting information about the interests and behavior of TV viewers collected in the Auditel (2003) and Eurisko (2002) studies about the Italian population. These studies enabled us to specify stereotypical preferences for several categories of TV programs that are coarser-grained than those of the General Ontology, but can be easily mapped to such categories (Gena, 2001). Thus, we specified a *Stereotype Ontology* defining the TV program categories to be considered and, similarly to the explicit preferences, we defined mapping rules that relate the corresponding user preferences.

The stereotypical descriptions include the specification of classification data and prediction information. This representation is similar to the one adopted in the SeTA system by Ardissono and Goy (2000). We sketch the representation by considering the stereotype describing the Housewife life style, shown in Figure 2.

Each classification datum is represented as a slot with three facets: the *Feature Name*, the *Importance* and the *Values*. The *Importance* describes the relevance of the feature

¹ The user may view and modify the form at any time.

to the description of the stereotype and takes values in [0,1]. The irrelevant features have importance equal to 0; the essential ones have importance equal to 1. The *Values* facet specifies a distribution of the feature values over the users represented by the stereotype. For each value, the percentage of individuals fitting it within the represented user class is specified. For instance, the interest in *Books* has medium importance in the characterization of the users belonging to the Housewife class (Importance is 0.6). Moreover, 80% of the housewives have low interest in reading books (frequency is 0.8).

| | |
|---|---|
| Housewife | |
| <u>Classification data</u> | |
| Age [<i>personal data</i>]: | Importance: 1, Values: (less_than_15, 0) (15/24, 0) (25/34, 0) (35/44, 0.5) (45/54, 0.5) (55/64, 0) (more_than_64, 0) |
| Gender [<i>personal data</i>] | Importance: 1, Values: (male, 0) (female, 1) |
| Books [<i>interest</i>]: | Importance: 0.6, Values: (low, 0.8) (medium, 0.2) (high, 0) |
| <u>Prediction part</u> | |
| movies-sentimental, Interest: 1; serial-soap, Interest: 1; TV news, Interest: 0.2; fashion programs, Interest: 0.5; cooking programs, Interest: 1; ... | |

Figure 2: The “Housewife” Stereotype

The slots in the *prediction part* of a stereotype describe the preferences of the typical user belonging to the represented class. In a prediction slot, the *Program category* specifies the described program category. Moreover, the *Interest* represents the user’s preference for the program category and takes decimal values in [0,1], where 0 denotes lack of interest and 1 is the maximum interest.

3.2.2. Management of the Stereotypical User Model

The user’s preferences are estimated in two steps. First, the user is matched against each stereotype *S* to evaluate how strictly her interests and socio-demographic data correspond to the interests and data of *S*. The result of this classification is a *degree of matching* with respect to each stereotype. This is a number in [0,1] where 1 denotes perfect match and 0 denotes mismatch.

In the second step, the user’s preferences are estimated by combining the predictions of each stereotype, proportionally to the degree of matching with the user. For each program category *C* of the Stereotype Ontology, the user’s interest in *C* is evaluated as the weighted sum of the interest predicted by the stereotypes; see Ardissono and

Goy (2000) and Ardissono et al. (2003) for details. Figure 3 shows the stereotypical user model of a user named Francesca.

| Interest in TV program categories | | Degrees of matching with stereotypes | |
|-----------------------------------|------|--------------------------------------|------|
| Movie-All | 0.65 | Colleagues | 0.34 |
| Movie-Sentimental | 0.65 | Engaged women | 0.24 |
| Movie-Comedy | 0.76 | Refined women | 0.26 |
| Movie-Detective | 0.46 | Dolphins | 0.16 |
| News All | 0.73 | ... | |
| Serial Fiction | 0.54 | | |
| ... | | | |

Figure 3: Portion of Francesca's Stereotypical User Model. The Predictions Have Confidence² = 0.43

3.2.3. Confidence in the Stereotypical Predictions

Having derived the stereotypes from broad studies such as the Eurisko one, we assume that the classes segment correctly the population of TV viewers. Thus, the confidence in the stereotypical predictions depends on the confidence that the user has been correctly classified by the system. In turn, this depends on the amount of information available at classification time and on "how stereotypical" is the user.

Confidence in the User Classification with Respect to a Stereotype

The confidence in the classification of the user in a stereotype S represents the confidence that the degree of matching is correct. This measure is evaluated by considering the minimum and maximum degrees of matching that the user might receive, if complete information about her were available.

- The lower bound of the degree of matching (DM_{min}) is evaluated by assuming that, for each classification datum the user has not specified, she matches the less frequent value of the datum, and by classifying her accordingly.
- The upper bound DM_{max} is evaluated by assuming that, for each missing classification datum, the user matches the most compatible value.

² This value derives from the confidence in the stereotypical classification and is the same for all the program categories because they are specified fully by the stereotypes. Other preferences, not shown in the figure, have lower confidence. Finally, the preferences not specified by the stereotypes have confidence equal to 0.

For instance, the lower bound of the compatibility of Age for “Housewife” is 0 and suits all the users younger than 35 or older than 55. The upper bound is 0.5 and suits the users between 35 and 54.

DM_{min} and DM_{max} define the interval of admissible values for the degree of matching (DM): $DM_{min} \leq DM \leq DM_{max}$. The larger is the interval, the lower the confidence in the classification has to be. In order to model this behavior, we have defined the confidence as:

$$conf_S = 1 - [(DM_{max} - DM_{min}) / \Delta]$$

Where Δ is the maximum distance between DM_{max} and DM_{min} . Δ is fixed for each stereotype and it corresponds to the case where no classification datum is set.

The formula defining the confidence in the user classification takes values in [0,1]. When the user is perfectly classified, $DM_{max} - DM_{min} = 0$ and $conf_S = 1$. When no information about the user is available $DM_{max} - DM_{min} = \Delta$ and $conf_S = 0$.

Confidence in the Predictions on the User’s Preferences

In order to evaluate the confidence in the predictions, an overall assessment of the quality of the user classification is needed that takes all the classes $\{S_1, \dots, S_n\}$ of the Stereotype KB into account. The average confidence in the user classification is an approximation of this measure:

$$Conf_{stereotypes} = (\sum_{i=1..n} conf_{S_i}) / n$$

However, this definition does not take the focus of the classification into account. As shown by our experiments (see Section 5), the most precise predictions are generated for the “very stereotypical” users matching a single stereotype or very few stereotypes. Moreover, low-quality predictions are generated for the users that match loosely several stereotypes. Thus, the confidence in the predictions is evaluated by combining the confidence in the classification ($Conf_{stereotypes}$ defined above) with an evaluation of its focus (*Focus*) in a fuzzy AND:

$$StereotypicalExpertConfidence = Conf_{stereotypes} * Focus$$

The focalization is derived from the evaluation of Shannon’s entropy on the degree of matching of the stereotypes. Suppose that $\{S_1, \dots, S_n\}$ receive $\{DM_1, \dots, DM_n\}$ values. Then, the entropy is evaluated as:

$$Entropy = \sum_{i=1..n} - DM_i * \log_2 DM_i$$

As the number of stereotypes is fixed, the entropy may be normalized in [0,1], therefore obtaining a normalized entropy *normEntropy*. The focalization is thus:

$$Focus = 1 - normEntropy$$

The focus takes the 0 value when the entropy is the highest, i.e., the classification is very uncertain. In contrast, when a single stereotype matches the user, the focalization is equal to 1. In turn, the confidence is only high when the classification relies on complete information about the user and is very focused.

3.3. THE DYNAMIC USER MODEL

3.3.1. Acquisition of Information about the User's Viewing Preferences

The Dynamic User Model specifies the user preferences for the program categories and sub-categories of the General Ontology and for the TV channels. As our system can track the user's actions on the TV, her viewing behavior can be explicitly related to the time of day when the actions occur. Thus, different from the other Experts, the preferences can be acquired for each viewing context and the user's habits during different weekdays and times of day can be identified.

In order to face the uncertainty in the interpretation of the user's viewing behavior, a probabilistic approach is adopted where discrete random variables encode two types of information: preferences and contexts. The sample space of the preference variables corresponds to the domain of objects on which the user holds preferences; the corresponding probability distributions represent a measure of such preferences (interests). The sample space of every context variable is the set of all the possible viewing times.

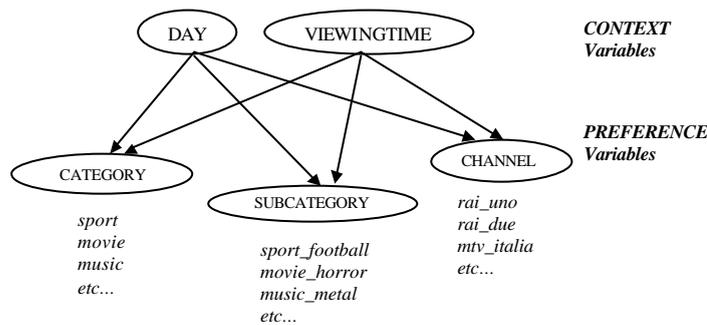


Figure 4: Portion of the BN that Represents the Dynamic User Model

Figure 4 shows the Bayesian Network (Neapolitan, 1990) used to represent the user preferences. In the network, the context variables are associated to the conditions in which the user preferences for the TV programs may occur. A context is characterized by temporal conditions represented by the *DAY* and *VIEWINGTIME* variables. These variables encode, respectively, 7 weekdays and 5 intervals of time in which the day can be subdivided. The context variables are root nodes in the network, since they are not influenced by any other information. The nodes of the Bayesian Network

(henceforth, BN) represent the user's contextual preferences, and they provide the probabilities for every program category, sub-category and channel.

For each user, the BN is initialized with a uniform distribution of probabilities on its nodes where all values assumed by the preference variables have equal probability. The BN is updated by feeding it with evidence about the user's selections of TV programs, starting from the first time she watches TV. Each time the user records a program, plays it³, or asks for more information about it, the system retrieves the category and the sub-category of the program and its transmission channel. Then, it feeds the BN with evidence that a new observation for that category is available.

The BN, implemented using the Norsys' Netica (2001) toolkit, predicts the user preferences by estimating the probabilities of different values for the category, sub-category and channel variables. Exploiting the values of the "DAY" and "VIEWINGTIME" variables generates the predictions.

Specifically, Netica provides a simple algorithm for parametric learning that takes the experience of each node of the BN into account. The experience of a node is defined as a function of the number of observed cases. The probability for the state node associated to a new observation is updated as follows:

$$new_prob = (prev_prob * prev_exper + learn_rate) / new_exper$$

where

- *learn_rate* is the learning rate of the observed action;
- *prev_prob* and *prev_exper* are the probability and the experience of the node, before the occurrence of the action;
- *new_exper* = (*prev_exper* + *learn_rate*) is derived from the previous experience by taking into account the learning rate of the observed action.

The probabilities of the state nodes associated to the types of actions that have not been observed are updated, for each viewing time, as follows:

$$new_prob = (prev_prob * prev_exper) / new_exper$$

Different learning rates are associated to the various action types in order to differentiate their impact on the learning phase. For instance, playing a TV program provides stronger evidence than asking for more information about it.

Figure 5 shows the viewing preferences acquired by observing the viewing behavior of user Francesca. The acquired preferences concern the Thursday-Evening context

³ The system tracks the time spent by the user on a program and compares it to the DVB specification of its duration.

and have been inferred by observing 60 performed actions: 30 Like, 10 Dislike, 3 Memo, 5 Record, 2 Play and 10 request of More Information.⁴

| Interest in TV program categories | | Interest in transmission channels | |
|-----------------------------------|-------|-----------------------------------|------|
| Movie-All | 1 | RAI 1 | 0.99 |
| Movie-Sentimental | 0.001 | RAI 2 | 0.63 |
| Movie-Comedy | 0.49 | Canale 5 | 1 |
| Movie-Detective | 0.3 | Telepiù Bianco | 0.56 |
| News All | 0.53 | Telepiù Grigio | 0.99 |
| Serial Fiction | 0.001 | ... | |
| Spot Politics / Society | 0.001 | | |
| ... | | | |

Figure 5 – Portion of Francesca’s Dynamic User Model (Evening Viewing Time).
The Confidence of the Predictions Is 0.5621765

3.3.2. Confidence in the Predictions of the Dynamic UM Expert

The confidence in the predictions is based on the quality of the data available to the BN. In turn, the quality depends on the amount of evidence about the user’s viewing behavior provided to the BN since the first time the user has interacted with the PPG. In fact, although some noise can be present in her behavior, the BN tolerates it in the presence of a large corpus of data. As the Dynamic User Model is initialized when no viewing data is available, the confidence must be initially equal to 0. The confidence may then increase as long as new user actions are captured by the system.

The Dynamic UM Expert computes the confidence in the predictions by counting how many user actions are observed for a specific context (experience of each node). A sigmoid function defines the confidence, given the number of observed actions. This function is normalized in the [0,1] interval and is defined below:

$$Conf(x) = 1/[1+e^{(k-x)*s}]$$

The function returns a confidence close to 0 if no action is observed in a specific context. Moreover, it returns a confidence of 0.5 after k actions are observed and the confidence gets close to 1 after the observation of $2*k$ actions. The s coefficient takes values in [0,1] and defines how steep the function has to be.

⁴ The interest values derive from the probability distributions computed by the BN. However, they are normalized in the [0,1] interval to be compatible with the interests predicted by the other UM Experts.

3.4. INTEGRATION OF THE PREDICTIONS PROVIDED BY THE UM EXPERTS

The predictions provided by the three Experts are combined by the UMC to estimate the user's preferences employed to personalize the recommendation of TV programs. The possibly conflicting predictions are reconciled by relying on their confidence and the result of this integration is stored in the Main User Model. More specifically, for each category P of the General Ontology, the predictions on P ($Interest_1, \dots, Interest_n$) provided by the Experts are combined into an overall *Interest* as follows:

$$Interest = \frac{\sum_{e=1}^n Conf_e * Interest_e}{\sum_{e=1}^n Conf_e}.$$

This formula merges the predictions in a weighted sum and normalizes the value in $[0,1]$ in order to let the most certain predictions influence the preference estimation in the strongest way. The confidence of the Experts may change along time; at least, the third Expert becomes more and more self-confident. Thus, their predictions are merged in different proportions and, eventually, the Dynamic UM Expert strongly influences the estimation of the user's preferences.

By integrating heterogeneous UM Experts we base the personalization on complementary types of information about the user. In fact, not all the user data are available during the same phases of the life cycle of the EPG. For instance, although the Dynamic UM Expert is expected to learn a precise user model, this module is not able to generate good predictions until a reasonable number of user actions are collected. Moreover, the Explicit Preferences Expert may be unable to provide predictions about several preferences because this specification is not mandatory⁵ (although the user may declare her preferences since the first interaction). Finally, the Stereotypical UM Expert may be unable to predict the user's interests if she does not provide her socio-demographic data, or if she clearly differs from stereotypical users. Figure 6 shows a portion of Francesca's Main User Model.

⁵ The data stored by this module may even be unreliable because the users are not always sincere. For instance, in the FACTS project (Bellifemine et al, 99), we noticed that the explicit preferences declared by users are often inconsistent with their real viewing behavior.

| Interest in TV program categories | | Interest in transmission channels | |
|-----------------------------------|------|-----------------------------------|------|
| Movie-All | 0.91 | RAI 1 | 1 |
| Movie-Sentimental | 0.28 | RAI 2 | 0.63 |
| Movie-Comedy | 0.49 | Canale 5 | 1 |
| Movie-Detective | 0.37 | Telepiù Bianco | 0.62 |
| News All | 0.79 | Telepiù Grigio | 1 |
| Serial Fiction | 0.58 | ... | |
| Spot Politics / Society | 0.42 | | |
| ... | | | |

Figure 6 – Portion of the Main User Model Describing Francesca’s Preferences for TV Program Categories in an Evening Viewing Time

4. Recommendation of TV Programs

The recommendation of TV programs is performed in two steps: first, the programs satisfying the user’s search query are retrieved and ranked with scores in the range [0,1] representing their suitability to the user. Then, the program list is sorted to reflect the user’s preferences and it is possibly pruned, if it includes too many items.

It should be noticed that the programs satisfying the user’s search query are retrieved from the system database of TV programs. This database is populated by downloading the program information from the satellite stream. The local storage of the TV content information is essential to support the generation of user-friendly EPGs because it enables the explicit representation of the relations between programs. For instance, the module responsible for populating the database unifies multiple occurrences of the same program, whenever possible⁶. Moreover, the module suitably classifies the serial programs. The availability of this type of information about programs supports the development of flexible presentation strategies. For example, our system simply presents the recommended programs by reporting all the occurrences of each program. However, summary recommendation lists could be generated by removing the redundancies; for example, the timing information of the same programs could be grouped.

⁶ The recognition of multiple occurrences is difficult when the information about the programs is delivered by different providers. In fact, although movies and serials are identified by their titles, different descriptions may be broadcast for other programs, such as sport events. The identification is anyway possible when the programs are broadcast by the same provider at different times because, in that case, the DVB information is consistent.

4.1. EVALUATING THE SCORE OF A TV PROGRAM

The generation of the scores for the individual TV programs is performed by considering both the user's interests in their program categories and her preferences for the transmission channels (preferences stored in the Main User Model). The former type of information represents the basis for the recommendations, instead we use the latter to refine the suggestions with evidence about the user's viewing habits at the different times of day. It should be noticed that the preference for the channel enables the system to take the user's preferences for individual programs into account without explicitly modeling the characteristics of such programs. In fact, the system relies on the criteria applied by the broadcasters in the selection of the programs to be shown. The scheduling of TV programs is based on the supposed TV audience in a given time slot that influences the quality and the characteristics of the programs.

The integration of the preferences for program categories and channels is performed according to the algorithm described below. Unfortunately, we cannot rely on complete information about TV programs because the fields of the DVB records broadcast in the satellite stream may be void. Therefore, more or less fine-grained preferences for program categories may be exploited to rank programs. If the program is classified in a sub-category of the General Ontology (e.g., the "Content" field of the descriptor is *sport_basket*), the corresponding user preference is employed. Otherwise (*sport*), the more general user preference is considered.

- (1) *Prog* = a TV program to be ranked;
- (2) *Cat* = category of *Prog* (retrieved from the descriptor of *Prog*);
- (3) *Ch* = transmission channel of *Prog* (retrieved from descriptor);
- (4) *Ctx* = current context (viewing time);
- (5) *Score* = user's interest in *Cat*, within *Ctx*;
- (6) *Interest_Ch* = user's preference for *Ch* in *Ctx*;
- (7) if *Interest_Ch* is significant
- (8) then *Score* = update *Score* according to *Interest_Ch*;

Given a TV program *Prog* to be ranked, the system retrieves the category of the program (2) and the transmission channel (3) from the descriptor. Moreover, the current viewing context, *Ctx*, is considered (4). Then, the system retrieves the user's preference (interest related to *Ctx*) for the program category in order to generate the first approximation of the score (5). Finally, the score is possibly refined (6-7-8) to take the user's preference (*Interest_Ch*) for the channel into account.

The approach adopted in the PPG relies on the following assumptions: no inferences can be made if the user's interest in the channel is medium. However, if the user

watches the channel very often at the time of day specified by Ctx , then this is positive evidence that she appreciates the programs usually broadcast at that time of day. Moreover, if she never watches the channel in a context Ctx , this is interpreted as moderate evidence that she does not like the programs broadcast by the channel at that time of day. Two relevance thresholds, set to 0.15 and 0.85, characterize the notions of low, medium and high interest for a channel. We have three cases:

1. *Medium preference for channel.* In this case $Score$ coincides with the user's preference for the Cat program category; no modification is performed. This happens when $Interest_Ch$, the interest in Ch during Ctx , is between 0.15 and 0.85.

2. *Very low preference for channel.* If the user's preference for Ch is very low ($Interest_Ch$ is between 0 and 0.15), the score of the TV event is decreased to represent the fact that the user typically does not watch Ch in context Ctx . Thus, the channel reduces evidence that she will like the specific program.⁷ In order to decrease the $Score$ proportionally with respect to the lack of evidence that the user watches the channel, but to maintain its value in $[0,1]$, $Score$ is updated as follows:

$$Score' = Score - \mathbf{a} * Interest_Ch * Score$$

Where \mathbf{a} , a decimal value in $[0,1]$, tunes the influence of the preference for the channel on the basic preference for the program category.

3. *Very high preference for channel.* If the user's preference for Ch is very high ($Interest_Ch$ is between 0.85 and 1), $Score$ is increased. In fact, $Interest_Ch$ provides positive evidence that the user likes watching the programs broadcast in Ch in the Ctx viewing time. In order to increase the $Score$ proportionally to the amount of positive evidence, but to maintain it in $[0,1]$, $Score$ is updated as follows:

$$Score' = Score + \mathbf{a} * Interest_Ch * (1 - Score)$$

Where \mathbf{a} is the same parameter used in case 2. In our experiments, \mathbf{a} is set to 0.1 to weakly influence the sorting strategy because we only want to change the order of programs belonging to the same category.

5. Experiments

5.1. THE EVALUATION METHODOLOGY

The recommendations of the PPG are generated by relying on the estimates of the user's preferences stored in the Main User Model (these preferences determine the

⁷ In some contexts, the user may not watch TV at all. Thus, the score of the programs is revised according to the channel preferences only during the viewing times where the user has medium or high preferences for at least one channel.

“space” devoted to the various TV program categories in the EPG). Thus, an evaluation of the system has to calculate the distance between the recommendations derived from these estimates and the real user’s preferences/selections. As the Main User Model results from the combination of the predictions of three UM Experts, we needed three kinds of information for a complete evaluation:

- a) The dataset exploited by the Stereotypical UM Expert to classify the users, i.e., socio-demographic data, general interests and lifestyles.
- b) The explicit users’ preferences for TV programs collected by the Explicit Preferences Expert.
- c) The users’ observed selections of TV programs, i.e., their viewing behavior.

To obtain this data we involved subjects belonging to the Auditel panel (Auditel, 2003). Auditel is the nonpartisan company that collects daily information about Italian TV audience. This survey classifies the Italian population in several socio-demographic panels according to the age, gender, education level, type of job and geographic zone. For each panel, the daily audience data is available, grouped by viewing time and TV channels. The Auditel panel includes 5.000 Italian families for a total number of 14.000 subjects. In order to collect datasets *a* and *b* we identified 62 Auditel subjects by following a non-probabilistic blocking sampling strategy. This is a sampling strategy that divides the population in layers related to the variables that have to be estimated, where each layer contains a number of individuals proportional to its distribution in the target population. We identified several layers characterized by different socio-demographic data, interests and TV program preferences. Every layer identifies a possible user of the PPG. We selected a small number of subjects because carrying out the required interviews and collecting the audience data was a complex task. Unfortunately, the complete analysis of the panel is not representative. However, we are currently extending our evaluation to other Auditel subjects to collect information about a representative sample of the Italian TV audience.

In order to acquire the previously mentioned data we operated as follows.

- We interviewed the subjects by means of a questionnaire. To obtain the desired information, we collected: general data (including personal data), information about general interests (books, music, sport, etc.), preferences for TV program categories and sub-categories. The final questionnaire included 35 questions where both the questions and the answers were fixed. The questionnaire was anonymous and introduced by means of a written presentation explaining the general research aims. For the items concerning the general data, the participants were required to check the appropriate answer out of a set of given answers. In the other questions, the subjects had to express their level of agreement with the options associated to the given questions by choosing an item in a 3-point Likert scale. The participants, without the presence of the interviewer, filled in the

questionnaires which were collected one week after the distribution. Then, we fed our PPG system with the acquired information to evaluate the validity of the user classification and the accuracy of the recommendations.

- After one year, we fed the PPG with the selections of TV programs made by the test subjects. This information was collected by the Auditel meter⁸ and stored in a database. In this way, we could activate the predictions generated by the Dynamic UM Expert. We entered the following information:
 - The context variables: day and viewing time.
 - The selected TV channel.
 - The (play) action suitably related to the TV program categories and sub-categories of our General Ontology.
 - The title of the watched program.

5.2. THE RESULTS

To separately test the recommendation capabilities of the three UM Experts we decided to evaluate the system's performance by simulating real scenarios where the Experts take different roles depending on the availability of information about the user. However, we did not evaluate the Explicit Preferences Expert that simply propagates the declared user preferences in the General Ontology.

We started from the Stereotypical UM Expert. We simulated an initial scenario where the user has specified her personal data and general interests, but where she does not declare her TV program preferences. Thus, the recommendations are based only on the stereotypical information. In this first phase, we evaluated the correctness of the stereotypical classification and the accuracy of recommendations by feeding the system with the socio-demographic data and the general interests (dataset *a*) collected by means of the interviews.

Then we simulated a scenario where the system has enough information (datasets *a*, *b* and *c*) to have the three Experts cooperating at the generation of the recommendations. In this second phase we also fed the system with the explicit program preferences and the TV program selections performed by the subjects.

5.2.1. Evaluation of the Stereotypical Classification

To evaluate the stereotypical classification, we compared the classification of the subjects computed by the PPG with the classification of two human Eurisko lifestyles

⁸ The meter is an electronic device connected to the TV that constantly monitors the viewing behaviour of the users belonging to the Auditel panel.

experts. The comparison showed that 70% of the users were classified correctly by the system. The remaining 30% were incorrectly classified for two reasons:

- The classification fails for “non-stereotypical” subjects, whose general interests differ from those evaluated according their socio-demographic data. Indeed, the Stereotypical UM Expert takes both socio-demographic and general interests into account to classify the user. However, the socio-demographic information plays a stronger role in the classification. Thus, if a user a has socio-demographic data typical of stereotype A , but her interests are typical of stereotype B , she is classified as belonging to stereotype A .
- The data provided by the Eurisko survey does not cover the whole Italian population. For instance the Retired stereotype only represents low-income users and the other retired users, such as the ex-managers, are not considered. This lack of information has to be overcome to improve the coverage of the stereotypical knowledge base and the consequent classification capabilities of the system.

The first issue deserves further discussion. The misclassification of a “non-stereotypical” user a causes wrong predictions because a prefers programs that would be recommended to the users belonging to another stereotype B . Indeed, we wanted to preserve the definition of the stereotypes and, at the same time, balance the contribution of the user’s socio-demographic data with that of her general interests. Thus, we employed the declared general interests as another source of information about the user’s preferences to be managed by the Explicit Preference Expert.

5.2.2. Evaluation of the System’s Recommendation Capabilities

We exploited the *Mean Absolute Error metric* (MAE^9) to evaluate the distance between the preferences predicted by the system and the users’ preferences/selections captured by monitoring their viewing behavior. Good et al. (1999) suggest that, in the evaluation of a recommender system, a satisfactory value of MAE should be about 0.7, in a range of 0.5. We also tested the accuracy of the recommendations by evaluating the *precision* of the collected data, i.e., the ratio between the user-relevant contents and the contents presented to the user; see Salton and McGill (1984).

First of all, we compared the TV program predictions generated by the Stereotypical UM Expert with the preferences expressed by the users and maintained by the Explicit Preferences Expert. Indeed, the users’ explicit preferences are expressed as

⁹ The MAE metric evaluates the distance between the system predictions and the user’s preferences/selections by means of rate vectors. A smaller value means more accurate predictions; see Good et al. (1999).

qualitative *low*, *medium* and *high* values. In order to compute the MAE by relying on similar measures, we exploited the numeric preference values generated by the Explicit Preferences Expert starting from the users' declarations. These values are reliable because the Expert derives them in a straightforward way from the qualitative ones.

For each test subject, after having entered the socio-demographic data, the general interests and the explicit preference values in the Explicit User Model, we calculated the differences between the values generated by the Explicit Preferences Expert and those generated by the Stereotypical UM Expert.¹⁰ Specifically, we evaluated the MAE by comparing the TV program category and sub-category predictions with the corresponding explicit preferences, with possible values ranging between 0 and 5. The obtained MAE value was 1,3 with precision 0,40; see Table 1.

Although this result cannot be considered satisfactory, we think that the MAE value was strongly influenced by the percentage of misclassified subjects, which was approximately 30%; see Section 5.2.1. Indeed, several subjects matched a high number of stereotypes. In these cases, the focalization of the classification was very low and downgraded the confidence of the predictions generated by the Stereotypical UM Expert, which were generic and corresponded to the users' real preferences in an approximated way. We think that these values might notably improve if we could extend the stereotypical knowledge base as described in Section 5.2.1.

| | |
|---|--------------------------|
| Stereotypical UM Expert | MAE = 1,3 Precision=0.40 |
| Stereotypical UM Expert + Explicit Pref Expert + Dynamic UM Expert | MAE = 0,3 Precision=0.80 |

Table 1. Evaluation of the System's Recommendations

In the second phase of our evaluation we compared the system's recommendation capabilities with the subjects' viewing behavior. We started from the information about the users already available to the system and we added the explicit preferences, which were omitted in the previous experiment. Next, we entered the TV program selections provided by the Auditel meter. More specifically, we fed the system with the observations collected during the first 10 months to train the Dynamic UM Expert. Then, we exploited those of the last 2 months to evaluate the distance between the system's recommendations and the subjects' observed selections. In this case, the

¹⁰ For the purpose of this evaluation, the recommendations generated by the three UM Experts are recorded in separated log files before being integrated in the Main User Model.

three Experts could generate reasonably confident recommendations and therefore an evaluation of the complete Main User Model was possible.

The resulting MAE was 0.30 and the precision was 0.80; see the second row of Table 1. These values are definitely satisfactory and confirm our hypothesis about the validity of the integration of different sources of information.

We also calculated an ANOVA to investigate the significance of the different MAE results obtained by considering the Stereotypical UM Expert alone and the final merge of the predictions provided by the three Experts. Our analysis showed that the different MAE results are due to a significant correlation between the Experts taken into account (independent variable) and the resulting program recommendations (dependent variable): $F(1,61) = 97,3$ $p < 0,01$.

6. Related work

Some recommender systems, such as MovieLens (2002), rely on collaborative filtering to personalize the suggestion of items. As discussed in Burke (2002), this technique performs well in domains where the set of items to be recommended is relatively stable, but has problems when new users, or new items, are considered. Other techniques, such as content-based filtering, support the recommendation of new items, but they tend to suggest items very similar to one another. In order to complement the advantages and disadvantages of different recommender systems, hybrid approaches are preferable in several application domains. Similar to the proposal described in Burke (2002), our PPG exploits different preference acquisition techniques, but the main difference is that we excluded collaborative filtering to focus on the techniques that can be efficiently applied locally to the user's set-top box.

The integration of the EPG in the set-top box is an important architectural feature of our system because it enables the continuous tracking of the user's viewing behavior. Thus, the user's preferences can be unobtrusively acquired while she watches TV, without requiring any explicit feedback.¹¹ In contrast, if a central server manages the EPG, the interaction with the TV is carried out in a distinct thread and can only be monitored while the user browses the program guide, unless special hardware is employed to connect the TV to the Internet. For instance, Smyth and Cotter's PTV Listings Service is based on a centralized architecture. To overcome the lack of connectivity with the TV, Smyth and Cotter propose the exploitation of GuideRemote, an interactive universal remote control that captures the user's selections while she watches TV.

¹¹ As a matter of fact, the user may rate programs, but the preference acquisition works well even without this type of information.

Another peculiarity of the PPG concerns the integration of alternative preference acquisition techniques to manage the hybrid user model. In other recommender systems, the most promising recommendation methods are selected by applying them in cascade (Burke, 2002), or by relying on a-contextual estimates of the precision of the methods or on the posterior evaluation of the recommendations. For instance, TvScout (Baudisch, 1998; Baudisch and Brueckner, 2002) combines a recommender based on the analysis of size-of-the-audience data in cascade with other two recommendation sources: the user's favorite program categories and the suggestions provided by *opinion leaders*, such as TV critics. As another example, in the TV Show Recommender (Zimmerman et al.) two implicit recommenders and an explicit one are fused by exploiting a neural network that tunes the influence of the competitors on the basis of the accuracy of their recommendations. Finally, the PTV Listings Service integrates a content-based recommender and a collaborative filtering one by merging the items best ranked by each recommender in a single suggestion list. Our approach differs from the previous ones in at least two aspects.

- On the one hand, our PPG combines heterogeneous inference techniques in a finer-grained way and clearly separates the estimation of the user's preferences from the generation of the personalized suggestions. We fuse three preference acquisition modules to acquire precise user models based on different information about the TV viewer. Then we put two recommendation techniques (content-based filtering and adjustment of rates based on the preferences for channels) in cascade to rank the TV programs.
- On the other hand, the system adopts a simpler approach to steer the fusion process. As described in Section 3.4, the PPG tunes the influence of the UM Experts in the estimation of the user's preferences by relying on the confidence in the predictions. Indeed, an accuracy measure should be coupled with the confidence one to evaluate the quality of the predictions in a more precise way. However, in the development of the PPG, we privileged the confidence measure, leaving the accuracy one for our future work, because the confidence can be exploited during the whole lifecycle of the EPG. In fact, it only depends on the amount of information about the user available to the Experts. In contrast, other accuracy measures take some time before being effective for new TV viewers.

The exploitation of stereotype-based techniques has a long tradition in the user modeling field, see Rich (1989); however, the definition of the stereotypical classes has been based on rather different assumptions about the population to be segmented. For instance, Kurapati and Gutta (2002) proposed to define stereotypical classes of TV viewers by clustering the viewing history data of a sample population. However, they noticed that some of the stereotypes created by the clustering algorithms did not make sense and were very difficult to understand. Instead, Barbieri et al. (2001)

proposed to define a set of classical stereotypes, such as Movie Lover and Film Freak, and let the TV viewer explicitly choose the one best matching her mood.

A deeper analysis of TV viewer stereotypes is proposed in a recent survey on the viewing preferences of Japanese TV viewers. Hara et al. group a sample of TV viewers on the basis of the features of the programs they say they have watched thoroughly. The results of the clustering analysis show that the viewers' interests influence their preferences for program categories; moreover, the people having the same socio-demographic attributes, such as age, gender and occupation, frequently differ in their preferences for TV program categories. Thus, Hara et al. propose viewing patterns as the most significant variable for the definition of stereotypical TV viewer classes. In particular, they define 8 viewer groups representing TV viewing "tastes" and watching styles, such as News/Culture Oriented, Diversion-Seeking Zapper, and so forth.

Although Hara et al.'s findings could discourage the exploitation of socio-demographic information about TV viewers to predict their preferences, we believe that the problem should be put in a different way. Specifically, socio-demographic information is not enough, but it is very useful when coupled with other information aimed at enriching the overall picture of the TV users. Indeed, the stereotypes exploited in our PPG are richer than the previously mentioned ones, as we derived them from complete studies of the TV viewer population under a *socio-demographic* and a *psychographic* point of view. In fact, the lifestyles survey we considered - Sinottica, conducted by Eurisko data analyzers (2002) - clusters the population in groups by taking into account not only socio-demographic data, but also consumer preferences, socio-cultural trends and homogeneous behaviors. Particularly, Sinottica is a psychographic survey on:

- Individuals (characteristics, values, behaviors, styles);
- What they consume (products/goods/ services and relative brands);
- Their exposure to the media (a survey in collaboration with Auditel, see 3.2).

By exploiting all these types of information, we could derive a set of stereotypes that partition the population in a precise way and reflect viewing preferences. Notice that these studies are exploited to plan the presentation of commercials within TV programs by the most representative content providers.

7. Conclusions and Future Work

This paper has presented the recommendation techniques applied in the Personal Program Guide (PPG). This is a prototype system generating personalized EPGs for set-top box environments. The PPG is based on a multi-agent architecture that

facilitates the integration of different user modeling techniques for the recognition of the TV viewer's preferences and the suggestion of the programs to watch.

As shown by our preliminary experimental results, the management of a hybrid user model, relying on different sources of information about the user's preferences, supports high-quality recommendations. This is not surprising: in fact, the recommendations based on explicit user information are subject to failures, because users are often unable to declare their real preferences. Moreover, the recommendations based on the observation of the user's viewing behavior take some time before being effective and, mirroring the user's usual selections, they fail to support the variety in the system's recommendations. In order to enable the system to generate high-quality suggestions since the first interaction with the user, we enriched the user models with community preferences by exploiting two main sources of information: on the one hand, the stereotypical preferences for program categories derived from lifestyle and audience data provide information about the preferences of similar TV viewers. On the other hand, the user's preferences for TV channels, at different times of day, support the refinement of the recommendations based on the audience analysis performed by the content providers.

In our future work, we want to extend the PPG in two main aspects. First, we want to enhance the TV program recommendations by taking household preferences into account and by refining the management of the hybrid user model. Second, we will redesign the User Interface to take usability issues into account.

Modeling household preferences is important because people rarely watch TV alone. As discussed in this volume by Masthoff, several recommendation strategies may be applied to satisfy the individual group members and avoid frustration. Although our system does not address household preferences, its recommendation capabilities can be extended in a rather straightforward way. In fact, the system architecture facilitates the integration of new User Modeling Experts and a Household Preference Expert could be added to handle group models. This UM Expert could employ the same preference acquisition techniques applied in our Dynamic Preference Expert to learn household profiles. Indeed, we believe that the most relevant issue to be solved is the automatic recognition of the user(s) in front of the TV. This issue is still unsolved, but some researchers, such as Goren-Bar and Glinansky (2002), are working to address it. The extension of the hybrid user model mainly involves the refinement of the fusion technique adopted to merge the predictions of the User Modeling Experts. As described in the previous part of this chapter, the predictions generated by the Experts are combined in a weighted sum, depending on their confidence. Although this approach has produced satisfactory results, we want to tune the fusion process by taking an accuracy measure into account, as well. In the multi-agent systems area, an established approach for the integration of possibly heterogeneous agents is based on

the joint evaluation of the agents' *self-confidence* and *reputation*. The self-confidence is a subjective evaluation of the agents' decision capabilities. The agent's reputation is an objective parameter evaluated by a third party. In the PPG, we already model the Experts' self-confidence that corresponds to the confidence in the preference predictions. Moreover, we will introduce the reputation that will be computed by the UMC by comparing the predictions provided by the Experts with the user's viewing behavior. The UMC will exploit the Experts' confidence and reputation to merge their preference predictions. Notice however that the UMC has to rely on the sole confidence for the fusion process until it has collected a significant amount of information about the user's viewing behavior.

As far as the User Interface is concerned, a lot of work has to be done to redesign it according to usability standards. However, as a first step in this direction, we want to focus on the presentation of the system's recommendations. At the current stage, the PPG suggests TV programs by coupling each item with a number of faces representing the recommendation degree. Moreover, the system limits the length of the recommendation list by omitting the presentation of the programs receiving very bad scores. As noticed in Zimmerman et al., the TV viewer's trust in the EPG would increase if she could be informed about all the available options, not only about the most interesting ones. The question is therefore how such possibly long list of alternatives could be presented in a clear and acceptable way, from the user's point of view. Other projects have encountered serious difficulties in making TV viewers accept prototype User Interfaces for Interactive TV; e.g., see Tinker et al. (2003). Therefore, some researchers are applying user-centered design to the definition of new User Interfaces for Electronic Program Guides; see van Barneveld and van Setten, in this volume.

Acknowledgments

The Personal Program Guide has been developed in a joint project between Telecom Italia Lab and the University of Torino. This work has been partially supported by the Italian M.I.U.R. (Ministero dell'Istruzione dell'Università e della Ricerca) through the Te.S.C.He.T. Project (Technology System for Cultural Heritage in Tourism). We are grateful to Flavio Portis, who helped us in the development of the Stereotypical UM Expert of the PPG.

REFERENCES

- Ardissono, L., Gena, C., Torasso, P., Bellifemine, F., Chiarotto, A., Difino, A., Negro, B. (2003). Personalized Recommendation of TV Programs. In *LNAI n. 2829. AI*IA 2003: Advances in Artificial Intelligence*, pages 474-486, Springer Verlag.
- Ardissono, L. and Goy, A. (2000). Tailoring the Interaction with Users in Web Stores. *User Modeling and User-Adapted Interaction* 10(4):251-303.
- Ardissono, L., Portis, F., Torasso, P., Bellifemine, F., Chiarotto, A., Difino, A. (2001). Architecture of a system for the generation of personalized Electronic Program Guides. *Proc. UM'01 Workshop on Personalization in Future TV*, Sonthofen, Germany.
- Auditel (2003). Auditel. <http://www.auditel.it>.
- Barbieri, M., Ceccarelli, M., Mekenkamp, G., Nesvadba, J. (2001). A personal TV receiver with storage and retrieval capabilities. In *Proc. UM'01 Workshop on Personalization in Future TV*, Sonthofen, Germany.
- Baudisch, P. (1998). Recommending TV programs on the Web: How far can we get at zero user effort? In *Recommender systems: papers from the 1998 workshop. Technical Report WS-98-08*, AAAI Press, pages 16-18, Menlo Park.
- Baudisch, P. and Brueckner, L. (2002). TV Scout. Lowering the entry barrier to personalized TV program recommendation. In *Proc. 2nd Int. Conf. On Adaptive Hypermedia and Adaptive Web-based Systems (AH2002)*, pages 58-68, Malaga, Spain.
- Bellifemine, F. et al. (1999). Deliverable A12D1 "Agent-system software for the AVEB Phase 1 Demonstrator". FACTS (FIPA Agent Communication Technologies and Services) is the ACTS project number AC317 of the European Commission.

- Burke, R. (2002). Hybrid Recommender Systems: Survey and Experiments. *User Modeling and User-Adapted Interaction* 12(4):331-370.
- Cotter, P. and Smyth, B. (2000). A Personalized Television Listing Service. *Communications of the ACM* 43(8):107-111.
- DVB (2000). Digital video broadcasting. <http://www.dvb.org>.
- Eurisko (2002). Sinottica. <http://www.eurisko.it>.
- Gena, C. (2001). Designing TV Viewer Stereotypes for an Electronic Program Guide. In *Proc. of the 8th Int. Conf. on User Modeling*, pages 274-276, Sonthofen, Germany.
- Good, N., Schafer, J.B., Konstan, J.A., Borchers, A., Sarwar, B.M., Herlocker, J.L., and Riedl, J. (1999). Combining collaborative filtering with personal agents for better recommendations. In *Proc. of the 16th National Conference on Artificial Intelligence*, pages 439-446, Orlando, Florida.
- Goren-Bar, D. and Glinansky, O. (2002). Family Stereotyping - A Model to Filter TV Programs for Multiple Viewers. In *Proc. AH'02 Workshop on Personalization in Future TV*, pages 101-108, Malaga, Spain.
- Hara, Y., Tomomune, Y., and Shigemori, M. (2004). Categorization of Japanese TV viewers based on program genres they watch. In *this volume*.
- Kurapati, K. and Gutta, S. (2002). TV Personalization through Stereotypes. In *Proc. AH'02 Workshop on Personalization in Future TV*, pages 109-118, Malaga, Spain.
- Masthoff, J. Group modeling: selecting a sequence of television items to suit a group of users. In *this volume*.
- Mostafa, J. (2002). *IEEE Intelligent Systems: Information Customization*, 17(6).
- MovieLens (2002). MovieLens: Helping you find the right movies. <http://www.movielens.umn.edu/>.

- Neapolitan, R.E. (1990). Probabilistic Reasoning in Expert Systems: Theory and algorithms. John Wiley & Sons, New York.
- Netica (2001). Application for Belief Networks and Influence Diagrams, User's Guide. <http://www.norsys.com>.
- Resnick, P., and Varian, H.R. (1997). *Communications of the ACM: Special Issue on Recommender Systems* 40(3).
- Rich E. (1989). Stereotypes and User Modeling. In Kobsa, A., Wahlster, W., editors, *User Models in Dialog Systems*, pages 31 -51, Springer Verlag, Berlin.
- Riecken, D. (2000). *Communications of the ACM: Special Issue on Personalization* 43(8).
- Salton, G. and McGill, M. (1984). Introduction to Modern Information Retrieval, McGraw-Hill Book Company.
- Smyth, B. and Cotter, P. The Evolution of the Personalized Electronic Programme Guide. In *this volume*.
- Tinker, P., Fox, J., Daily, M. (2003). A zooming, electronic programming interface. 3rd Workshop on Personalization in Future TV (TV'03), pages 7-11, Johnstown, PA.
- TiVo, Inc. (2002). TiVo: TV your way. <http://www.tivo.com>
- van Barneveld, J. and van Setten, M. Designing Usable Interfaces for TV Recommender Systems. In *this volume*.
- Zimmerman, J., Kurapati, K., Buczak, A.L., Schaffer, D., Martino, J., and Gutta, S. TV Personalization System. In *this volume*.