Personalisation Technologies for the Digital TV World

Paul Cotter and **Barry Smyth** [1]

Abstract. The information overload problem is becoming a serious issue on the Internet and users are finding it increasingly difficult to quickly locate the right information at the right time. Content personalisation techniques may provide a useful solution by enabling information services to respond to the implicit and explicit preferences of individual users. In this paper we describe an innovative personalised TV listings service (PTV), which has been successfully deployed on a number of Irish Web sites and which has attracted over 15,000 users in the first year of operation.

1 INTRODUCTION

Today the World-Wide Web contains over 800 million pages of information accessed by more than 200 million users. Many users experience difficulties in locating the right information at the right time - this is the information overload problem. Indeed this problem is set to worsen as a new generation of users access the Web using mobile-phone like devices and the Wireless Application Protocol (WAP) - the luxury of a large screen as a viewing portal onto the Web is not yet a possibility and instead users must cope with a screen that is a fraction of the size of even a modest monitor.

One approach to solving the information overload problem involves the development of personalisation technologies so that information systems can automatically learn about the information needs of individual users so that content could be *personalised* to match the needs of individuals; for example, irrelevant content could be withheld and relevant content highlighted [1, 3, 7, 9].

This paper describes the PTV family of Internet systems for personalising TV listings content - each user receives a TV guide that has been customised for their viewing preferences [11]. PTV has been licensed to, and deployed on, a number of Irish Web sites since early 1999 (PTV at www.ptv.ie and MyTV at www.ireland.com). We will focus on MyTV, licensed to the Irish Times newspaper group, and also introduce a new version of PTV designed to deliver personalised TV guides to mobile phone users.

2 PROBLEM DESCRIPTION

The latest digital TV systems promise viewers an unprecedented level of programme choice; tens of TV channels today will become hundreds tomorrow, and thousands soon after. Even now, many subscribers have access to upwards of one hundred channels, broadcasting over 2500 programmes per day. These changes will introduce a new information overload problem as it becomes increasingly difficult to find out what programmes are on in a given day or week. This signals an end to the paper-based TV guide - consider a 300 page publication for each week's viewing!

The digital TV vendors are aware of these issues, and do recognise the start of a serious information overload problem. A current solution is the Electronic Programme Guide (EPG), providing users with on-screen access to online TV listings (see Figure 1). However, simply providing an electronic equivalent of the paper TV guide is not a scalable solution to the problem. For example, the EPG shown in Figure 1 covers a 90-minute time-slot for only 7 channels. This means that even a relatively modest line-up of 70 channels will occupy 10 screens of information for each 90-minute slot, or 160 screens for each viewing day.

Figure 1. A modern EPG (courtesy of ReplayTV).

Some EPG's try to help the user further by compiling genre-based guides, based on comedy or drama of films, for example. However, this option is still relatively limited, and at best provides only short-term relief from the information overload problem, after all, there may still be hundreds of comedies showing on a given night, and many of these may be of no interest to a given user.

Any real solution to this information overload problem requires an understanding of the viewing preferences of users to enable EPG's to adapt information for individuals, filtering out irrelevant programme content, and transforming the space of viewing options from an intractable cast of thousands to a manageable few.

3 APPLICATION DESCRIPTION

The PTV project is motivated by the belief that techniques such as user profiling [5, 6], machine learning [2], and content recommendation [1, 3, 8, 10] hold the key to a new generation of intelligent,

[1] Smart Media Institute, Department of Computer Science, University College Dublin, Belfield, Dublin 4, Ireland, email: {Paul.Cotter, Barry.Smyth}@ucd.ie

ECAI 2000. *14th European Conference on Artificial Intelligence*
Edited by Werner Horn
Published in 2000 by IOS Press

personalised EPGs. In this section we describe the PTV system, focusing on its core personalisation components. An architectural overview and system demonstration is also provided.

3.1 The PTV Architecture

PTV is a Java-based client-server system and includes a specially designed optimised, multi-threaded server and dynamic HTML/WML page generator, plus all of the artificial intelligence and user profiling components necessary for content personalisation. The basic system architecture is shown in Figure 2.

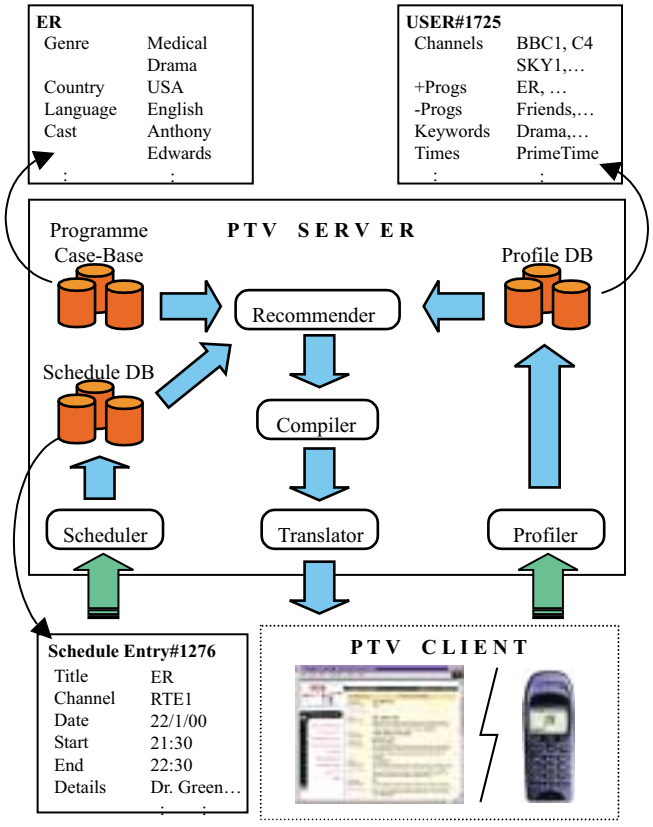

Figure 2. The PTV Architecture.

Profiler: This agent tracks and logs the activities of users as they request and grade personalised guides. Each guide will contain a list of programme recommendations for a user, and each user is encouraged to grade these recommendations (see Section 3.2) - this information is translated into programme, channel, viewing time, and genre preferences. The profile agent represents all of this information in a secure profile database (see Figure 2) and each user can view and edit their own profile as they see fit.

Recommender: The recommender agent is the core of PTV. Its job is to select a list of relevant programmes for a particular user based on their user profile. The recommender uses a hybrid recommendation technique combining case-based and collaborative filtering recommendation approaches (see Section 3.2). We argue that the combination of these techniques ensures an accurate and robust recommendation agent.

Compiler: Once a list of relevant programmes has been compiled for a particular user, their broadcast details (channel, time, description, video+ code, etc.) are pulled from the schedule database to produce a formatted guide for that user.

Translator: The translator converts the generic guide format produced by the compiler into the appropriate format for the target display device, HTML or WML (Wireless Markup Language).

Scheduler: This agent maintains the schedule database by accessing online TV listing resources. The scheduler collects programming information for up to a day in advance to offer PTV users viewing information for today's and tomorrow's programmes (see Figure 2 for an example schedule entry).

3.2 Core AI Techniques

In this section we describe PTV's user profiling and content recommendation techniques driving the personalisation process.

3.2.1 User Profiling

A user profile is made up of *domain preferences and programme preferences*. The former captures general user preferences such as a list of available TV channels, preferred viewing times, genre preferences, and guide format preferences. Programme preferences are represented as two lists of graded programme titles, a positive list containing programmes that the user has liked in the past, and a negative list containing programmes that the user has disliked. Each list also contains grading information to indicate the strength of the user's like or dislike for a particular programme.

PTV gathers profile information in two ways. First, users can manually edit their profile (see Figure 3). This is probably the most reliable way to update a user's profile but it places a burden on the user, and our studies suggest that while users are happy to provide initial information about their viewing preferences (at registration time, see Section 3.3.1), they are less likely to manually edit their profile later on (current statistics show that only 12% of updates are of this kind).

The second way to collect profile information is by allowing users to grade recommendations (see the grading icons in Figure 4); this accounts for about 88% of the profile updates. The profiler uses grading information to automatically alter a user's profile by updating the programme preference lists (adding positively or negatively graded programmes) and by adapting domain preferences such as channel and viewing time preferences. This long-term feedback connection between user and system is vital if PTV is to maintain an accurate picture of each user over time.

3.2.2 Case-Based Recommendation

Content-based recommendation techniques have their roots in the Information Retrieval (IR) community. The basic idea is to recommend content items that are similar to other items that a user has liked in the past [1, 4, 11]. A variety of techniques can be used to determine content similarity (eg, keyword matching). Case-based recommendation follows a similar strategy, but bases its similarity on the type of similarity functions and structured content representations found in many case-based reasoning systems [4, 12].

In PTV each programme is encoded as a feature-based description stored in the case-base; features such as genre, actors, and country of origin are used (see Figure 2). The similarity between a given programme and a user profile is actually based on the similarity between the programme case and the profile schema. The profile schema is a feature-based encoding of a user profile that is compatible with the programme cases. The similarity can then be computed using

the standard weighted-sum similarity metric as shown in equation 1; Where $f_i^{Schema(u)}$ and f_i^p are the i^{th} features of the schema and the programme case respectively.

$$PrgSim(Schema(u), c) = \sum_i w_i \bullet sim(f_i^{Schema(u)}, f_i^c) \quad (1)$$

Content-based methods may require significant knowledge-engineering effort to develop suitable representations and similarity models, and this is especially true for case-based methods, which demand structured content representations. PTV at the moment relies on a manual approach to updating the programme case-base.

Furthermore, because content-based methods make recommendations based on item similarity, the newly recommended items tend to be similar to the past items leading to reduced diversity. In the TV domain this can result in narrow recommendation lists; for example, a profile containing many preferred comedies will often lead to the recommendation of other comedies, and the minor user preferences can become swamped.

3.2.3 Collaborative Recommendation

Collaborative recommendation methods such as automated collaborated filtering are an alternative to content-based techniques. Instead of recommending new items that are similar to the ones that the user has liked in the past, they recommend items that other *similar users* have liked [1, 2, 3, 7, 8, 10]. Instead of computing the similarity between items, we compute the similarity between users, or more precisely the similarity between user profiles. In PTV the recommendations for a target user are based on the viewing preferences of the k most similar users.

$$PrfSim(u, u') = \frac{\sum_{p(u) \cup p(u')} |r(p_i^u) - r(p_i^{u'})|}{4 \bullet |p(u) \cup p(u')|} \quad (2)$$

$$PrgRank(p, u) = \sum_{u' \in U} Pr\ fSim(u, u') \quad (3)$$

PTV computes user similarity by using a simple graded difference metric shown in equation 2; where p(u) and p(u') are the ranked programmes in each user's profile, and $r(p_i^u)$ is the rank of programme p_i in profile u. The possible grades range from -2 to +2 and missing programmes are given a default grade of 0 (for alternative similarity techniques see [2, 10]).

Once PTV has selected k similar profiles for a given target user, a recommendation list is formed from the programmes in these similar profiles that are absent from the target profile. This list is then ranked and the top r programmes are selected for recommendation. The ranking metric is shown in equation 3; U is the subset of k nearest profiles to the target that contain a programme p. This metric biases programmes according to their frequency in the similar profiles and the similarity of their recommending user. In this way popular programmes that are suggested by very similar users tend to be recommended.

Collaborative filtering solves many of the problems associated with content-based methods. For example, there is no need for content descriptions or sophisticated similarity metrics. Moreover, recommendation diversity is maintained as relevant items that are dissimilar to the items in a user profile can be suggested.

Collaborative filtering does suffer from two shortcomings. There is a startup cost associated with gathering enough profile information to make accurate user similarity measurements. There is also a latency problem since new items will not be recommended until they have found their way into sufficiently many profiles. This is problematic in the TV domain because new and one-off programmes occur regularly and do need to be considered for recommendation once they become available.

3.2.4 A Hybrid Recommendation Strategy

The key to PTV's success is its use of a hybrid recommendation approach that combines the case-based and collaborative techniques. Both techniques operate independently during recommendation but their results are pooled to produce the final recommendation list. Each guide will therefore contain a selection of recommendations, some from the content-based technique and others from collaborative recommendation technique.

This results in a recommendation scheme that combines the strengths of content-based and collaborative methods while at the same time eliminating their major individual disadvantages. For example, new and one-off programmes (ignored by the collaborative recommendation technique) can be recommended by the content-based method, while the collaborative technique will ensure that recommendation diversity is maintained, and will also ensure that programmes that are missing from the case-base can be considered for recommendation.

3.3 System Demonstration

This section demonstrates the MyTV system, a re-branded version of PTV licensed to the Irish Times newspaper group as part of their ireland.com portal site. MyTV has been in operation since August 1999. It has attracted 8000 registered users in the first 6 months and delivers over 30,000 personalised guides per month to Irish TV viewers.

3.3.1 Registration and Preferences

Each user must register with PTV to benefit from the personalisation facilities so that they can be recognised and tracked (using cookie techniques) each time they use the system. After a new user has chosen a username and password, the main registration task is to provide some initial TV preferences to bootstrap the recommendation process. The user can select channel preferences, viewing times, genre keywords, and programme preferences as shown in Figure 3.

3.3.2 Personalised Guides

Once registered, a user can access the personalised guides. The main interface has been kept simple with all site features and guide options available from a simple menu. The user can receive personalised guides for programmes on today or tomorrow, as well as what's on now and next, a list of the top 10 programmes (compiled from the current user profiles), a wide range of genre-based guides (comedy, drama, film, etc), plus full TV listings.

Figure 4 shows the main personalised guide for today's TV (22/1/2000). In total the guide contains 15 programmes selected on 10 different channels. The portion shown in Figure 4 includes three programmes that the user is known to like (The Simpsons, Star Trek, and Brookside) and three new programmes that have been selected as relevant (Technofile, Casualty, and ER). Importantly, beside each of these new programmes is a set of grading icons (thumbs-up and

Figure 3. The MyTV preferences page allows users to edit their profiles.

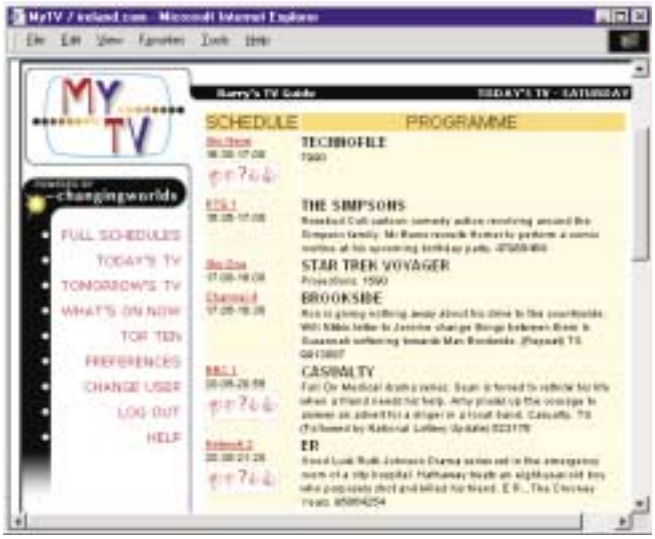

Figure 4. The main "Today's TV" personalised guide.

thumbs-down icons) to allow the user to rate each of these new recommendations. A five-point scale from bad to good is used, and in this case all three recommendations are on-point; the user is interested in technology programmes (Technofile) and has more than a passing interest in medical dramas (Casualty and ER).

PTV (and MyTV) provides a number of other guide types. For example, a second type of guide contains a full listing for a given channel (in terms of time and title) but with all relevant programmes highlighted with additional content. Thus, the benefits of personalisation are still available even when the user wishes to view a complete channel lineup.

PTV also compiles *themed* guides - that is, guides that are personalised with respect to a particular subject category (comedy, drama, music, movies, news and sport, etc.). These guides are analogous to the genre-based guides offered by some modern EPGs. They are automatically produced by PTV's recommendation engine by creating profiles to represent *virtual* users with an appropriate set of subject specific viewing preferences.

4 APPLICATION BUILDING

The PTV project began in 1997 as part of a basic research programme in the Department of Computer Science at University College Dublin. The project team comprised of a single PhD student and a supervisor. The PTV research prototype was completed in late 1998 and during the next 6 months developed as the final PTV system. In total an estimated 40 months of person effort had been invested in this final system.

This final system is a Java-based client-server system and runs on Linux on an Intel 450MHz processor with 64MB of RAM. It has been stress-tested beyond 7 million hits per month without any substantial performance degradation. Moreover, because this system has been developed with commercial licensing options in mind, it can be readily customised to meet client requirements. For example, the MyTV version, produced for the Irish Times electronic publishing division, required approximately 6 person weeks of effort and involved integration with existing client databases storing TV listings content and user information.

From a maintenance viewpoint PTV requires only minimal resources. The daily update of TV listings content is automatic, and the only time that real maintenance is required is when a new set of channels need to be added to the system. This can be managed easily by a support engineer with minimal experience.

5 APPLICATION BENEFITS

PTV's central objective is to provide a solution to the information overload problem associated with TV listings, and whether or not PTV is successful will depend on its ability to produce accurate TV guides within an easy to use system. To test this a comprehensive user evaluation was performed during 1999. A total of 310 users (both new and experienced) completed a detailed questionnaire and the important results are summarised in Figure 5.

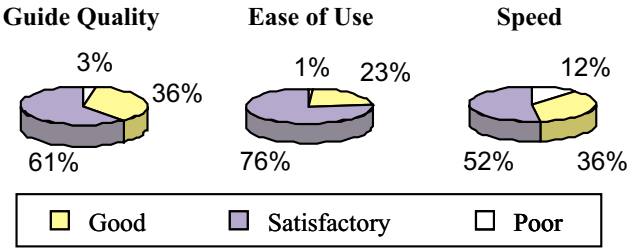

Figure 5. Summary user evaluation results.

The results are clearly very positive with the majority of users satisfied with PTV in terms of its personalisation quality, ease of use,

and speed. Notably, only 3% faulted the personalisation quality, and only 1% faulted the system's ease of use. Moreover, even though 12% of users critised PTV's speed, this has to be viewed positively given the limited speed of today's Internet and the fact that PTV is dynamically generating personalised guides.

PTV's users are not the only beneficiaries of its personalisation features. For example, the MyTV version of PTV provides an important service as part of the ireland.com portal site and has so far attracted over 8000 new regular users (requesting more than 30,000 personalised guides per month). Thus, MyTV provides a valuable "sticky" content service, driving increased traffic through the ireland.com site, and helping to build loyalty among its users

6 CONCLUSIONS & FUTURE WORK

We have motivated, described, and demonstrated PTV, a novel Internet system for delivering personalised TV listings content. The system has been licensed to, and deployed on, a number of Irish Web sites since early 1999 with considerable success; to date the PTV systems (PTV and MyTV) have attracted more that 15,000 users and generate over 50,000 personalised guides per month.

We believe that PTV's personalisation technology can form the core of the next generation of WAP services. We have adapted the PTV system for WAP and plan to launch this new system in the near future. Figure 6 shows an example of a personalised guide from the WAP-PTV system. Notice that because of presentation space and memory restrictions the structure and format of the guide has had to be significantly altered (compared to MyTV). The single-guide Web format, containing programme title, channel, time, details, and grading icons, has been split into multiple pages. The main guide contains programme titles only, with extra pages used to hold the channel and time information for each programme, and a further page for the programme details and grading options.

We are currently working on using the PTV technology to develop a range of personalised information services in a variety of other information domains and we believe that similar successes will be forthcoming for these future systems.

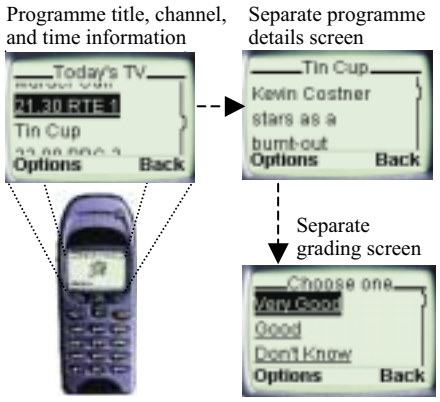

Figure 6. The WAP-PTV system adapts PTV for use on the new generation of WAP devices. The screen shots illustrate portions of a WAP-PTV guide.

REFERENCES

[1] M. Balabanovic and Y. Shoham, 'Fab: Content-based collaborative recommender.', *Communications of the ACM*, **40**(3), 66–72, (1997).

[2] D. Billsus and M.J. Pazzani, 'Learning collaborative Information Filters', in *Proceedings of the International Conference on Machine Learning*, (1998).

[3] D. Goldberg, D. Nichols, B.M. Oki, and D. Terry, 'Using collaborative filtering to weave an information tapestry', *Communications of the ACM*, **35**(12), 61–70, (1992).

[4] K.J Hammond, R. Burke, and K. Schmitt, *A Case-Based Approach to Knowledge Navigation*, chapter A Case-Based Approach to Knowledge Navigation, 125–136, MIT Press, 1996.

[5] A. Jennings and H. Higuchi, 'A user model neural network for a personal news service', *User Modeling and User-Adapted Information*, **3**(1), 1–25, (1993).

[6] J. Kay, 'Vive la Difference! Individualised Interaction with Users', in *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, ed., Chris Mellish, pp. 978–984. Morgan Kaufmann, (1995).

[7] J.A. Konstan, B.N. Miller, D. Maltz, J.L. Herlocker, L.R. Gordan, and J. Riedl, 'Grouplens: Applying collaborative filtering to usenet news', *Communications of the ACM*, **40**(3), 77–87, (1997).

[8] D. Maltz and K. Ehrlich, 'Pointing the Way: Active Collaborative Filtering', in *Proceedings of the ACM Conference on Human Factors in Computing Systems* , pp. 202–209. ACM Press, (1995).

[9] M. Perkowitz and O. Etzioni, 'Adaptive Web Sites: An AI Challenge. ', in *Proceedings of the 15th International Joint Conference on Artificial Intelligence*, (1997).

[10] U. Shardanand and P. Maes, 'Social Information Filtering:Algorithms for Automating 'Word of Mouth'', in *Proceedings of the ACM Conference on Human Factors in Computing Systems* , pp. 210–217. ACM Press, (1995).

[11] B. Smyth and P. Cotter, 'Surfing the Digital Wave: Generating Personalised TV Listings using Collaborative, Case-Based Recommendation', in *Case-Based Reasoning Research and Development. Lecture Notes in Artificial Intelligence*, eds., Klaus Dieter Althoff, Ralph Bergmann, and L.Karl Branting, pp. 561–571. Springer Verlag, (1999).

[12] I. Watson, *Applying Case-Based Reasoning: Techniques for Enterprise Systems*, Morgan-Kaufmann, 1997.