PTV: Intelligent Personalised TV Guides

Paul Cotter & Barry Smyth

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

Abstract

Although today's world offers us unprecedented access to greater and greater amounts of electronic information, we are faced with significant problems when it comes to finding the right information at the right time – this is the essence of the *information overload problem*. One of the proposed solutions to this problem is to develop technologies for automatically learning about the implicit and explicit preferences of individual users in order to customise and personalise the search for relevant information. In this paper we describe the development of the PTV system (Personalised Television Listings – http://www.ptv.ie) which tackles the information overload problem associated with modern TV listings data, by providing an Internet-based personalised TV listings service so that each registered user receives a daily TV guide that has been specially compiled to suit their particular viewing preferences.

Introduction

The term *information overload* has become almost synonymous with the Internet and the World Wide Web, and today the Internet's 200+ million users are finding it increasingly difficult to efficiently locate precisely relevant information content among its growing repository of 500+ million pages. For example, modern search engines provide only a first cut through the information space, leaving the user with a significant search task in order to locate individual information items. This is beginning to cause problems on the Internet and is seen as a serious barrier to its future success.

This problem takes on even more significance when one considers the new generation of mobile phones, which offer users an alternative Internet access route through the Wireless Application Protocol (WAP). Web content (including text, graphics, forms, hyperlinks etc.) is displayed as WML encoded pages; WML is the equivalent of HTML on WAP devices. At the present time, these devices suffer from greatly reduced display sizes, limited bandwidth, and restricted on-board memory (see Figure 1). Under these conditions it becomes even more important to be able to offer WAP users personalised information content, since current WAP devices do not facilitate a trawl through even moderate quantities of information in conventional Web terms.

Figure 1. Compared to the traditional mode of Web access (that is, the PC), WAP-based access devices, such as mobile phones, can suffer from greatly reduced screen real-estate, bandwidth, and page sizes.

Content personalisation is one potential solution to the information overload problem. It promises the precise delivery of user-targeted information by automatically learning about the preferences of individual users over time, and by using this information to guide the search for, and presentation of, relevant information.

In this paper we focus on an emerging information overload problem that is associated with the new generation of digital TV systems. We suggest that it will become

almost impossible for people to cope with the promise of hundreds of TV channels and thousands of TV programmes daily, and that traditional TV guides will fail to provide any practical assistance. We present PTV (http://www.ptv.ie) as a real solution to this problem. In short, PTV is an innovative Internet service that uses content personalisation techniques to automatically learn about the TV viewing preferences of individual users to provide them with highly customised and personalised daily TV guides. In particular, we focus on two versions of PTV, a fully deployed Web-based system and a newly developed version for WAP-based mobile phones (making PTV one of the world's first fully personalised WAP information services for the mobile phone market).

Problem Description

With the arrival of new cable and satellite television services, and the next generation of digital TV systems, we will soon be faced with an unprecedented level of programme choice. Tens of TV channels today will become hundreds of channels tomorrow, and thousands soon after that. Even today, in Europe and the US, many subscribers have access to upwards of one hundred channels, broadcasting over 2500 programmes per day. The service providers tell us that this new level of channel choice will revolutionise the way we use and view TV, but they rarely tell us of the pit-falls that lie just around the corner. These developments will introduce a whole new set of information overload issues since we have not yet developed the tools that are necessary to deal with this new level of choice. It will become increasingly difficult to find out what programmes are on in a given week, never mind locating a small set of relevant programmes for a quiet evening's viewing.

Figure 2. An example electronic programme guide (EPG) courtesy of BskyB.

Consider the traditional TV guide, which lists programming information for perhaps a week in advance. The days of a slim, easy to digest 30-page volume are numbered. Instead we are faced with TV guides of telephone book proportions, running into hundreds of pages of indigestible schedule charts. Moreover, the way that we interact with our TV sets will also have to change. For better or for worse, "channel surfing" will become a thing of the past as a means for finding out what's on now – while a rapid surf through 20 or 30 channels takes an acceptable few minutes, surfing through 200 or 300 channels could take a number of hours.

Of course the digital TV vendors are aware of such issues, and do recognise the beginnings of a serious information overload problem. They are now offering electronic programme guides (EPGs) to help users to navigate through the TV listings maze. These guides provide an on-screen menu system for searching online TV listings information. Figure 2 shows an example of Sky Digital's EPG, listing programmes on 10 channels for a 1 hour time-slot. However, Sky's full 60 channel line-up requires up to 6 screens of information for each viewing hour (that is, over 140 screens per viewing day). Clearly the burden of search remains with the user and these EPGs face the same problems of scale as existing TV guides.

Some EPG's attempt to help the user further by providing a genre-based view of the listings data. For example, a user might request a list of all comedies, or drams, or films that are on a given day, and this will help to focus the search further. However, these static genre-based approaches are still relatively crude, and at best provide 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.

Application Description

The PTV project is motivated by the belief that the TV listings domain can benefit greatly from an EPG that incorporates content personalisation techniques as a means of filtering and customising TV listings information for individual users (Kay 1995; Perkowitz and Etzioni 1997). In this section we describe the PTV system, focusing in particular on how it produces personalised TV guides by integrating user profiling, case-based reasoning, and collaborative profiling techniques.

Hardware & Software

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 personalisation. It currently runs on WindowsNT on an Intel 450MHz processor with 64MB of RAM and has been stress-tested beyond 7 million hits per month without any substantial performance degradation.

System Architecture

PTV users can register, login, and view their personalised TV guides as specially customised HTML pages (for conventional PC-based access) or as WML pages (for mobile phone access). The architecture of PTV (Figure 3) does not depend on the mode of access (PC vs. WAP-based device) and all user interaction is handled via HTTP. The heart of the system lies with its server-side components,

which handle all the main information processing functions such as user registration and authentication, user profiling, guide compilation, and the all-important programme recommendation and grading.

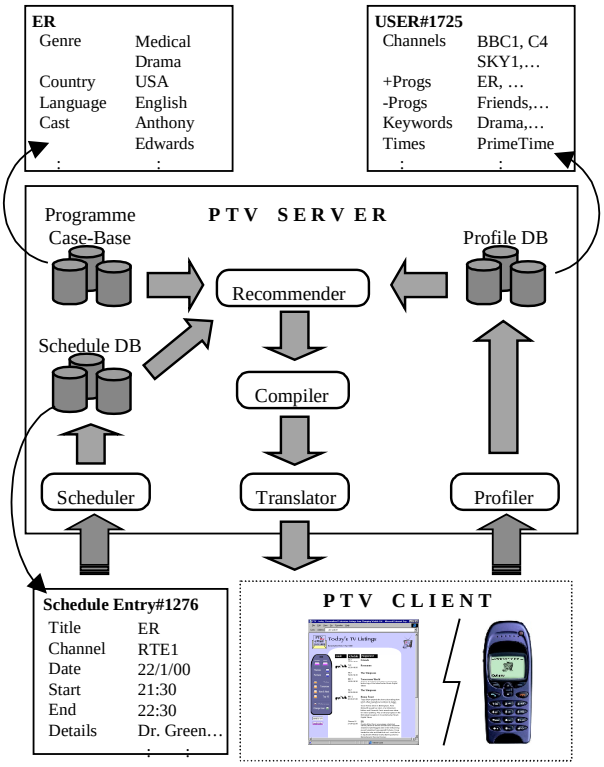


Figure 3. An overview of the PTV system architecture.

Profile Database & Profiler: The key to PTV's personalisation facility is an accurate database of user profiles. Each profile encodes the TV preferences of a given user, listing channel information, preferred viewing times, programme and genre preferences, guide preferences etc. (see Figure 3). Preliminary profile information is collected from the user at registration time in order to bootstrap the personalisation process. However, the majority of information is learned from grading feedback provided by the user; each recommended programme is accompanied with grading icons or links that allow the user to explicitly evaluate the proposed recommendation (see also Section 4.1).

Programme Case-Base: This database contains the programme content descriptions (programme cases). Each entry describes a particular programme using features such as the programme title, genre information, the creator and director, cast or presenters, the country of origin, and the language; an example programme case for the comedy 'Friends' is shown in Figure 3. This information repository is crucial for the content-based (case-based) recommendation component of PTV (see Section 4.2).

Schedule Database: This database contains TV listings for all supported channels. Each listing entry includes details such as the programme name, the viewing channel, the start and end time, and typically some text describing the programme in question (see the schedule entry example in Figure 3). The schedule database is constructed automatically from electronic schedule resources.

Recommender: The recommender component is the intelligent core of PTV. Its job is to take user profile information and to select new programmes for recommendation to a user. In the next section we will explain how PTV uses a hybrid recommendation approach that combines content-based and collaborative recommendation strategies (see Sections 4.2 and 4.3).

Guide Compiler: To compile a personalised guide for a user, PTV uses two programme lists: (1) programmes listed as positive in the user's profile, along with those programmes selected for recommendation (that do not occur in the profile); (2) a list of programmes to be aired on the specified date by channels listed in the user's profile. The intersection of these lists is the set of programmes that will finally appear in the personalised guide.

Guide Translator: The guide compiler produces a generic guide format, which is automatically converted into a HTML or WML page by the guide translator as appropriate. While individual guides are converted into single HTML pages for the Web, they are converted into multiple WML pages (or cards) for mobile phone usage; this is necessary to solve the problems of limited presentation space (and memory space) that exist on current WAP phones

Problem Description

Artificial Intelligence techniques are central to the success of the PTV system. Specifically, the ability to accurately personalise the television guide of an individual user relies on the availability of an accurate model of this user (user profiling), and an ability to relate this profile to relevant programme content (programme recommendation). In this section we outline PTV's user profiling component and its content recommendation strategies

Acquiring User Profiles

The success of PTV depends ultimately on the quality of its personalised guides, and this depends largely on the quality of the user profiles and their ability to represent the viewing preferences of users (Jennings and Higuchi 1993; Kay 1995; Perkowitz and Etzioni 1997)). In PTV each user profile contains two types of information: *domain preferences* and *programme preferences*. The former describe general user preferences such as a list of available TV channels, preferred viewing times, subject keywords and genre preferences, and guide format preferences. Programme preferences are represented as two lists of 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.

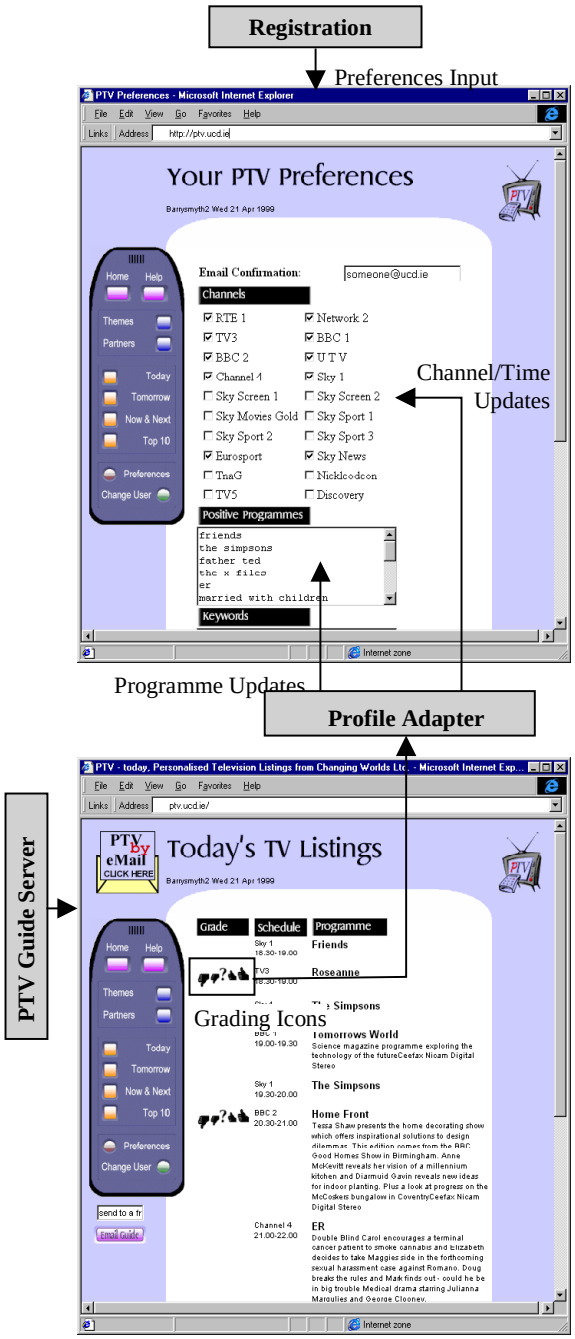


Figure 4. User profiles and feedback in PTV (Web-based).

Profile information is gathered in two ways. Users are encouraged to manually update their profiles directly by specifying viewing preferences. However, while manual profile editing has its advantages (usually in terms of profile accuracy) it is a burden for the users. In particular, we have found that users are happy to provide fairly complete domain preferences but tend to provide only limited programme preferences. For this reason, PTV includes a profile update facility that is driven by direct user feedback through a set of grading icons listed beside guide programmes. PTV's profiler uses this information to automatically alter a user's profile in a number of ways. The simplest modification is to update the programme preference lists by adding positively or negatively graded programmes to the appropriate list. However, the domain preferences can also be altered. For example, viewing time preferences can be adjusted if a user frequently prefers prime-time programmes to morning shows. In general, this long-term feedback connection between user and system is vital if PTV is to maintain an accurate picture of each user over time.

Figure 4 outlines how user profiles are updated in the Web-based PTV system. A similar scenario operates in the WAP-based version of PTV except that preference and grading options require a number of individual pages, rather that having a single preferences page or integrating the grading icons with the main guide pages as in the Web-based version (see Section 5 for further details and example screen shots).

A Content-Based Recommendation Approach

Ultimately in PTV, personalising a given user's TV guide boils down to recommending the right programmes for that user given their various viewing constraints. PTV harnesses two complementary recommendation strategies to base its recommendations on the programmes that a given user has liked in the past (case-based or content-based) and on the programmes that similar users like (collaborative). In this section we look at the more traditional content-based (or case-based) approach (Watson 1997) and in the following section we will look at the complementary collaborative recommendation strategy.

The basic philosophy in content-based recommendation is to recommend items that are similar to those items that the user has liked in the past; see also (Balabanovic and Shoham 1997; Hammond et al. 1996; Smyth and Cotter 1999). For PTV, this means recommending programmes that are similar to the programmes in the positive programme list and dissimilar to those in the negative programme list. Three components are needed for content-based recommendation: (1) content descriptions for all TV programmes (see the programme case-base in Section 2 and Figure 3); (2) a compatible content description of each user's profile; (3) a procedure for measuring the similarity between a programme and a user.

PTV's programme case-base has already been outlined (Section 3.2) and an example case is shown in Figure 3. Each case is described as a set of features and the similarity between two cases can be defined as the weight sum of the similarity between corresponding case features. However, there is no direct means of computing the similarity between a case and a user profile, as user profiles are not described as a set of case features. Instead each raw user profile is converted into a feature-based representation called a *profile schema*. Basically, the profile schema corresponds to a content summary of the programme preferences contained in a user profile, encoded in the same features as the programme cases. The similarity between a

profile and a given programme case 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.

$$1.\ \mathrm{PrgSim}(\mathrm{Schema}(u), p) = \sum w_i \cdot \mathrm{sim}\left(f_i^{\mathrm{Schema}(u)}, f_i^{p}\right)$$

A problem with content-based methods is the knowledge-engineering effort required to develop case representations and similarity models. 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 lot of comedies if the majority of profile programmes are comedies.

A Collaborative Recommendation Approach

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 (Balabanovic and Shoham 1997; Billsus and Pazzani 1998; Goldberg et al. 1992; Konstan et al. 1997; Maltz and Ehrlich 1995; Shardanand and Maes 1995). And 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.

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. Of course this is just one possible similarity technique that has proved useful in PTV, and any number of techniques could have been used, for example statistical correlation techniques such as Pearson's correlation coefficient (see eg., Billsus and Pazzani 1998).

$$2.\ \mathrm{PrfSim}(u, u') = \frac{\sum_{p(u)\ p(u')} \left| r\left(p_i^u\right) - r\left(p_i^{u'}\right) \right|}{4 \cdot \left| p(u) \quad p(u') \right|}$$

$$3.\ \mathrm{PrgRank}(p, u) = \sum_{u' \ U} \mathrm{PrfSim}(u, u')$$

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 suggested by very similar users tend to be recommended.

Collaborative filtering is a powerful technique that solves many of the problems associated with content-based methods. For example, there is no need for content descriptions or sophisticated case similarity metrics. In fact, high quality recommendations, that would ordinarily demand a rich content representation, are possible. 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 some shortcomings. There is a startup cost associated with gathering enough profile information to make accurate user similarity measurements. There is also a latency problem in that new items will not be recommended until these items have found their way into sufficiently many user profiles. This is particularly problematic in the TV domain because new and one-off programmes occur regularly and do need to be considered for recommendation even though these programmes will not have made it into any user profiles.

The key to PTV's success is the use of a combined recommendation approach. For a given guide, a selection of programmes is suggested, some are content-based recommendations (including new or one-off programmes) while others are collaborative recommendations. In particular, recommendation diversity is ensured through the use of collaborative filtering and the latency problem can be solved by using content-based methods to recommend new or one-off programmes.

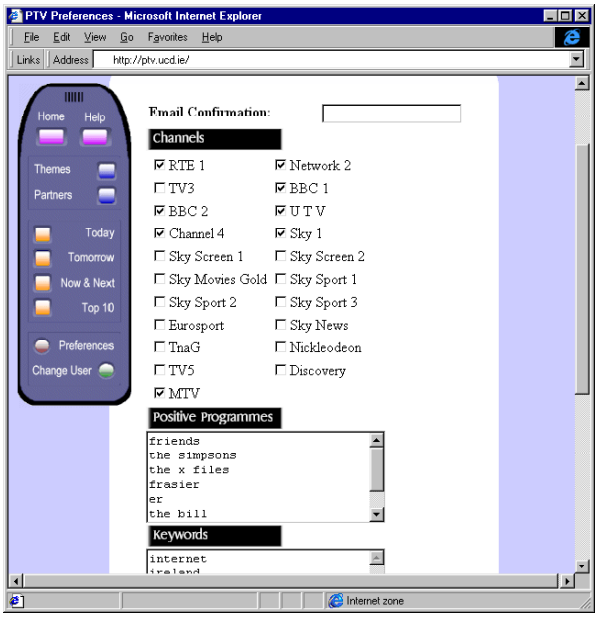
System Demonstration

In this section we look at the use of the PTV system by a new user, stepping through each of the basic stages, from initial registration through to guide viewing. To avail of PTV's personalisation facilities, each new user must register an account with PTV. In addition to submitting the usual username and password details the user is also asked to provide initial profile information (as discussed in Section 4.2) and this is shown by the screen shots in Figure 5. Figure 5(a) shows the preferences screen for the web-based version of PTV, whereas Figure 5(b) shows the equivalent screens in the WAP-based version.

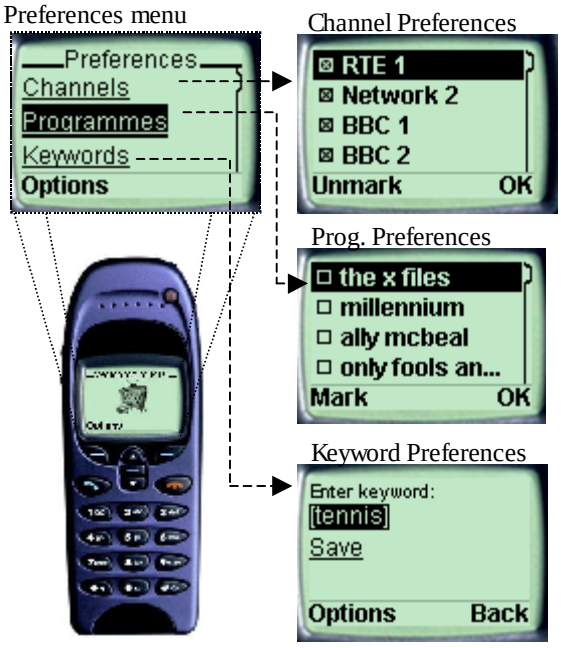
The presentation restrictions introduced by the current generation of WAP enabled mobile phone should now be clear. As a result the single-screen preferences of the Web-based version is replaced by multiple screens in the WAP-based version.

Incidentally, access to the preferences information shown in Figure 5 is not restricted to registration time. Users can access and edit their profiles at any time as a means of supplementing any updates made by the programme grading process.

Combined channel, prog., and keyword preferences



(a) Web-based Preferences Screen



(b) WAP-based Preferences Screens

Figure 5. The initial profile screen during user registration.

Combined prog., channel, time, and grading information

(a) Web-based PTV Guide

(b) WAP-based PTV Guide

Figure 6. A personalised guide for today's TV.

Once a user has registered they can access their personalised guides. The main interface has been kept simple and all of the site features and guide options are available from a simple menu on both the Web and WAP versions. 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 (as compiled from the user profiles currently in the system), a wide range of subject-specific guides (eg, comedy, drama, film, etc), plus of course full TV listings.

Figure 6(a&b) shows examples of a personalised daily guide from the Web and WAP versions of PTV. On the Web-based version, Figure 6(a), we can see four programmes from BBC, Channel 4 and UTV; the full guide contains about 10 programmes. Two of the visible programmes have been selected for this guide because the user is known to enjoy them (Friends and The Bill are both listed in the user's profile). Two of the programmes (Eastenders and Auntie's Sporting Bloomers) are *recommendations* based on this user's profile, and as such they are annotated with grading icons so that the user can provide feedback on the quality of these recommendations. Similar information is carried by the WAP version but because of presentation and memory restrictions it is not possible to display complete guide information (programme name, details, channel, time, and grading options) on a single screen. Instead, the main guide screen contains only programme title, channel, and time information as shown. A user is free to request more information on any particular programme by a simple selection, and at this point can continue on to grade that programme as desired; see Figure 6(b).

PTV's personalisation facility is also used to compile a variety of so-called *themed guides* which are analogous to the genre-based guides used by modern EPGs (see Section 2). These guides are not compiled with respect to a particular user but rather with respect to a particular theme. For example, PTV provides themed guides for comedy, sport, news, drama, soap operas, chat shows etc, and the programmes for these guides are automatically selected using PTV existing recommendation engine by creating profiles to represent 'virtual' users with specialized interests in a particular theme.

Deployment, Evaluation, and Maintenance

PTV was originally developed as part of a three-year basic research project in the Department of Computer Science at University College Dublin, Ireland. The resulting personalisation technology was re-implemented as a commercially viable personalisation engine during 1999 (approximately 9 person months). A well developed set of tools and systems now exist for rapidly developing new commercial versions of PTV to suit a wide range of client needs. For example, the latest development of the WAP-based version of PTV required only 8 person-weeks of development time.

Application Use and Payoff

PTV went live in January 1999 and the number of registered users has grown to nearly 20,000 with about 50,000 personalised guides generated per month. Furthermore, as yet PTV has not been publicised, so its current popularity is based largely on word of mouth and unsolicited press coverage. In fact, since the launch the site has received a Yahoo! Site of the Month award plus many favourable press reviews from Irish and UK magazines and newspapers including ComputerActive, Dot.ie, PC Live, Business & Finance, the Sunday Business Post, and the Irish Times. PTV was a finalist in the Irish Golden Spider Internet Awards under the best use of technology category.

PTV has been so successful that a campus company called Changing Worlds (http://www.changingworlds.com) has been established in University College Dublin to market the PTV system and its underlying personalisation technology. For example, a re-branded version of PTV has already been licensed and launched on the Ireland.com portal site (http://www.ireland.com) run by the Irish Times newspaper group. This re-branded version is called MyTV and since its launch in August 1999 has attracted approximately 8,000 new users to produce approximately 40,000 personalised guides per month. MyTV adds a valuable 'sticky' content service to the ireland.com site and helps to secure user loyalty and increase traffic levels.

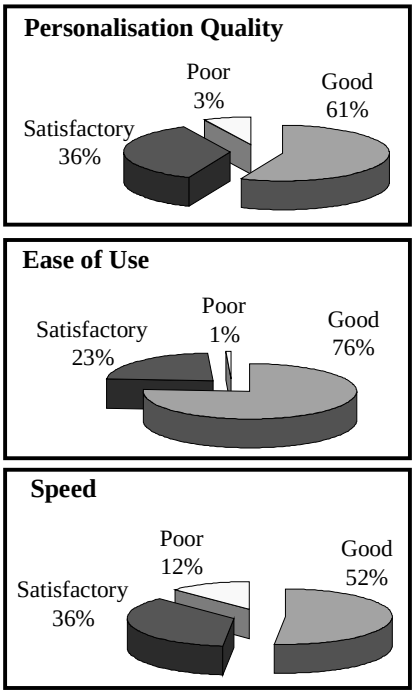


Figure 7. User survey summary results.

End-User Evaluation

Of course, from a scientific viewpoint, the big question concerns the accuracy of PTV's personalisation facility.

From January to June, 1999, real users carried out an extensive and detailed study on PTV. In total 310 users completed a comprehensive questionnaire regarding all aspects of PTV including its personalisation quality, speed, and ease of use, the results of which are summarised in the pie charts shown in Figure 7; see Smyth & Cotter (1999) for further experimental results.

Clearly, the results are extremely positive. Critically, only 3% of users found the personalised guides to be of poor quality and 99% of users found the site to be easy to use as a source of TV listings. Moreover, 88% of users found the response time to be acceptable, which we view very positively, especially considering the fact that PTV's pages are created dynamically and the limited speed of today's Internet.

Maintenance

The PTV systems are designed to have minimal ongoing maintenance requirements. For example the maintenance of the user profiles and the programme schedules is fully automatic. In fact, in our experience, the only manual maintenance that is required involves the addition of new programme cases and the addition of new channels and cable-regions. Even a relatively inexperienced user can manage both of these maintenance options by using PTV's in-built tools.

Conclusions

We believe that PTV represents a convergence of technologies that provides an effective solution to the very real problem of providing people with relevant TV listings information as digital TV becomes a reality. PTV personalises TV information to meet the viewing preferences of individual users by integrating two different information-filtering strategies, case-based reasoning and collaborative filtering, with user profiling techniques. The resulting hybrid personalisation technique allows programme recommendations to be made according to the type of programmes a target user has enjoyed in the past as well as the programmes that other similar users have enjoyed.

To date this technology has been deployed on the Internet as the PTV (www.ptv.ie) and MyTV (www.ireland.com) web sites. These applications have proven to be successful with widespread adoption across Irish Internet market. In the near future, a WAP version of PTV will be formally launched and we believe a similar success story will unfold as mobile phone users recognise the real benefits of high-quality content personalisation on their restricted mobile handsets. In fact, we argue that traditional TV listings services are not appropriate given the screen and bandwidth limitations of the current generation of WAP devices – a personalised service such as PTV is the best available solution.

Of course PTV's personalisation technology is not restricted to personalising TV listings content. The PTV systems are built around a content personalisation engine that can be readily adapted to practically any source of information content and Changing Worlds is currently using this technology to develop the next generation of intelligent, personalised information services.

References

Balabanovic M., Shoham Y. 1997. FAB: Content-Based Collaborative Recommender. *Communications of the ACM*,40(3): 66-72.

Billsus, D. & Pazzani, M. J. 1998. Learning collaborative Information Filters. In *Proceedings of the International Conference on Machine Learning*, Wisconsin, USA.

Goldberg D., Nichols D., Oki B. M., Terry D. 1992. Using Collaborative Filtering to Weave an Information Tapestry. *Communications of the AC* 35(12) 61-70.

Hammond, K. J., Burke, R., and Schmitt, K. 1996. A Case-Based Approach to Knowledge Navigation. In *Case-Based Reasoning Experiences Lessons and Future Directions*. MIT Press, 125-136

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

Kay J. 1995. Vive la Difference! Individualised Interaction with Users. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, 978-984. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence Inc.

Konstan J. A., Miller B. N., Maltz D., Herlocker J. L., Gordan L. R., Riedl J. 1997. Grouplens: Applying Collaborative Filtering to Usenet News. *Communications of the ACM*. 40(3) 77-87.

Maltz D., Ehrlich K. 1995. Pointing the Way: Active Collaborative Filtering. In *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '95)*. ACM Press, New York, N.Y., 202-209.

Perkowitz, M. & Etzioni, O. 1997. Adaptive Web Sites: An AI Challenge. In *Proceedings of the 15th International Joint Conference on Artificial Intelligence*. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence Inc.

Shardanand, U. & Maes, 1995. P. Social Information Filtering: Algorithms for Automating 'Word of Mouth'. *In Proceedings of the Conference on Human Factors in Computing Systems (CHI95).* ACM Press, New York, N.Y. 210-217

Smyth, B. & Cotter, P. 1999. Surfing the Digital Wave: Generating Personalised TV Listings using Collaborative, Case-Based Recommendation. In *Proceedings of the International Conference on Case-Based Reasoning*, 561-571. Springer-Verlag, Germany.

Watson, I. (1997) *Applying Case-Based Reasoning: Techniques for Enterprise Systems*. Morgan-Kaufmann.