TV Content Recommender System

Srinivas Gutta, Kaushal Kurapati, KP Lee, Jacquelyn Martino, John Milanski, J. David Schaffer, John Zimmerman

Philips Research
345 Scarborough Rd.
Briarcliff Manor, NY 10510
{srinivas.gutta, kaushal.kurapati, kp.lee, jacquelyn.martino, john.milanski, dave.schaffer, john.zimmerman}@philips.com

Abstract

The plethora of content available to the consumer has become overwhelming. Increasing amounts of information are being disseminated through terrestrial broadcast, satellite, and cable leading to an information overload. Common modes of searching for TV programs currently in existence include: TV-guide, PreVue channel and rudimentary search tools available through satellite dish TV programming service. These tools are general-purpose in nature and are not specifically tailored to the individual viewer's taste. Towards that end we advance in this paper a recommender system that searches for TV programs based on their likes/dislikes through implicit personalization techniques.

Introduction

Today, most consumers face an exhausting task of having to find something to watch on TV that fits their interests. Currently, the main modes for searching and identification of relevant TV content are the following: (a) browsing through pages of the 2D grid-format paper TV guide, (b) waiting on the PreVue channel, which is an automated scrolling version of the paper TV-guide grid and (c) using the rudimentary search tools that are available with a satellite dish TV programming service. The specific problems associated with them include the inability of the viewers to weed out irrelevant content, non-interactivity and difficulty in navigation. An alternative way is to provide viewers with a personalized means to provide intuitive user interfaces and the ability to filter program information. We propose a combination of the following two approaches to assist the users:

Search Engines & Information Visualization - Provide them with easy-to-use tools for search and present the abstract information in a way that is intuitive and easy to comprehend.

Recommender Systems - Provide them with a system that tracks and recognizes their preferences and organizes the TV program content accordingly.

Background

Researchers in information visualization have taken several approaches to visualize multidimensional data in various application contexts [Card *et al*., 1999]: finding movies and homes, news articles in a database, and web pages on the Internet. The specific problems associated with them include the use of scatter plots which are not particularly helpful in visualizing text-based TV program data, due to occlusion problems and the use of immersive visualization techniques leading to slower response time of the system, and loss of user orientation. The literature is rich with descriptions of other visualization systems for scientific and data mining applications; however, researchers have not focussed their attention on the consumer domain, which is what our work addresses.

The research in adaptive systems aims to build recommender systems that help the user in filtering information based on the user's profile. Several such systems have been built in recent years to help users deal with various sources of information [Etzioni, 1999]. However, these systems have a major source of information missing, the TV. The TV Advisor by Das and Horst [Das and Horst, 1998] is one example of a recommender system for TV found in the literature. They make use of explicit techniques to generate recommendations for a TV viewer. Such techniques require the user to take the initiative and explicitly specify their interests, in order to get high quality recommendations. Implicit techniques, on the other hand, provide a non-intrusive approach to lessen the burden on the user by inferring the user's preferences from the use of a TV set.

Application Prototype

Our application prototype aims to be a 'smart electronic program guide (smart EPG)' that enables a user to search and browse through a TV programs database. It is 'smart' because it maintains an adaptive user profile and makes recommendations of TV programs, computed according to the profile. The application can be divided into 3 parts as described below.

The search environment provides the tools for the user to formulate a search for retrieval from a TV-programs database. We organize the valid search criteria along 'strings' or 'bracelets', which represent individual dimensions of the multi-dimensional TV programs database. Each 'bracelet' is a grouping of complementary 'beads' where each 'bead' is a visual representation of information contained in the database. In our prototype, we have 7 bracelet categories: day of the week, time of the day, program genres, channels, keywords, user profile names and saved searches. This notion of beads evolves from ancient prayer beads or the abacus, where the beads served as information holding units and were useful for counting. The navigation and selection of search criteria takes place through a standard remote control.

The overview environment is concerned with the visual representation of the search results. The results are TV shows matching the search criteria, retrieved from the database. Our approach to visualization was to map this abstract information in a manner that parallels the human tendency to put physical objects closer when they are important and to let them fade into the background as their importance decreases. We use depth as a cue to achieve this notion. A tunnel model was used for displaying the results comprising of rings, each of which serves as a placeholder for the recommended TV shows. The TV-shows that are highly recommended for a user will be displayed closer to the user, on the first few rings, than those that are not.

In the current version of our prototype, user profiles can be used as search criteria to generate system recommendations of TV content. The search results are visualized, in the overview environment, in the order of high to low relevancy based on a desirability score that is computed according to the person's profile. The scores are the output of either of two recommender engines that we have built as part of our research. Both of our approaches to generating TV program recommendations are based on implicit profiles of TV viewers. At the current time, we pursued approaches that could deal with incremental updates of the viewing history, and inconsistencies in the indexing of shows in the TV show data. An implicit profile is built from the viewing history of a TV viewer. The viewing history is a list of shows that a viewer has watched (positive examples) and not watched (negative examples). The implicit nature of our profiling method stems from the fact that the process does not involve any explicit interaction with TV viewers, regarding their likes and dislikes, other than collecting information about what shows have been watched.

The first of our TV program recommender systems uses the Bayesian classifier [Billsus and Pazzani, 1996] approach to compute the likelihood that the viewer will like or dislike a particular TV program. We approach the problem with a 2-class Bayesian decision model, where a show belongs to either the class, watched, or the class, not watched. The user profile, in the Bayesian context, is a collection of attributes (or features) together with a count of how many times an attribute occurs in positive and negative examples. From this profile, we first compute the prior probability that a show belongs to a particular class and then the conditional probability that a given feature will be present if a show is in either of the two classes. Using these probabilities we finally compute the a posteriori probability for a new show, given its feature set, that it belongs to a particular class.

The second approach construct rules for classifying shows given a *training set* of positive and negative shows that are part of the TV viewing history. We begin by deriving a decision tree (DT) which is then decomposed into rules for classifying the shows. The decision tree employed is Quinlan's C4.5 [Quinlan, 1993] that uses an information-theoretic approach based on entropy. C4.5 builds the decision tree using a top-down, divide-and-conquer approach: it first selects an attribute, then divides the training set into subsets characterized by the possible values of the attribute, and follows the same procedure recursively with each subset until no subset contains objects from more than one class. The single-class subsets correspond to the leaves of the decision tree, while a node indicates that a further test needs to be performed on that show to determine which class the show belongs to. When a new show, which is not part of the training set, is encountered, the DT is parsed to obtain a probabilistic class distribution for the show and the class with the highest probability is the predicted class.

Future Work

We have tested our application with potential users in order to get an appraisal of our visualization and personalization techniques. We got very encouraging endorsements of our concepts. Users also gave helpful suggestions for improvement, which are under consideration.

References

[Billsus and Pazzani, 1996] D. Billsus and M. Pazzani. Revising User Profiles: The Search for Interesting Web Sites. *Proceedings of 3rd International Workshop on Multistrategy Learning*. AAAI Press, 1996.

[Card *et al*., 1999] Stuart Card, Jock Mackinlay, and Ben Shneiderman. *Readings in Information Visualization.* Morgan Kaufmann Publishers Inc., San Francisco, California, 1999.

[Das and Horst, 1998] Duco Das and Herman ter Horst. Recommender Systems for TV. *Proceedings of 15th AAAI Conference*, Madison, Wisconsin, July 1998.

[Etzioni, 1999] Oren Etzioni, Jorg Muller, Jeffrey Bradshaw. Proceedings of the 3rd Annual Conference on Autonomous Agents, ACM Press, 1999.

[Quinlan, 1993] Ross J. Quinlan. C4.5*: Programs for Machine Learning*. Morgan kaufmann Publishers Inc., San Francisco, California, 1993.