

3PRS: a personalized popular program recommendation system for digital TV for P2P social networks

Jui-Hung Chang · Chin-Feng Lai · Yueh-Min Huang ·
Han-Chieh Chao

Published online: 7 November 2009
© Springer Science + Business Media, LLC 2009

Abstract Digital TV channels require users to spend more time to choose their favorite TV programs. Electronic Program Guides (EPG) cannot be used to find popular TV programs. Hence, this paper proposes a personalized Digital Video Broadcasting — Terrestrial(DVB-T) Digital TV program recommendation system for P2P social networks. From the DVB-T signal, we obtain EPG of TV programs. The frequency and duration of the programs that users have watched are used to extract programs that users are interested in. The information is collected and weighted by Information Retrieval (IR). The program information is then clustered by k-means. Clusters of users are also grouped by k-means to find cluster relationships. In each group, we decide the most popular program in the group according to the program weight of the channel. When a new user begins to watch the TV program, the K-Nearest Neighbor (kNN) classification method is used to determine the user's predicted cluster label. Then, our system recommends popular programs in the predicted cluster and similar clusters.

Keywords Digital video broadcasting-terrestrial (DVB-T) · Electronic program guide (EPG) · Information retrieval (IR) · K-nearest neighbor (kNN) · k-means

J.-H. Chang (✉) · C.-F. Lai · Y.-M. Huang
Department of Engineering Science, National Cheng Kung University, Tainan, Taiwan 701,
People's Republic of China
e-mail: changrh@mail.ncku.edu.tw

C.-F. Lai
e-mail: cinfo@www.mmn.es.ncku.edu.tw

Y.-M. Huang
e-mail: huang@mail.ncku.edu.tw

H.-C. Chao
Institute of Computer Science & Information Engineering and Department of Electronic Engineering,
National Ilan University, I-Lan, Taiwan 260, People's Republic of China
e-mail: hcc@mail.niu.edu.tw

H.-C. Chao
Department of Electrical Engineering, National Dong Hwa University, Hualien, Taiwan 974,
People's Republic of China

1 Introduction

Electronic Program Guide (EPG) devices, which consist of a and a digital set-top box [12], give users information about programs. EPGs usually have a graphical interface which shows the program name, channel, and broadcast time. Users can see detailed information about each program; but in some cases, a EPG may only provide a text-based interface for users [5]. Users can search for programs according to topic, quickly select a program, and use parental control functions. Due to the increase of the number of TV programs, recommendation systems have been introduced [22, 23]. TV-Scout and PTV are web-based recommendation systems; P-EPG and Multi-Agent are STB-based recommendation systems [8, 21]. WAP-based recommendation systems are also available (such as PTV). The recommendation mechanisms of these systems are similar; Information retrieval and information filtering methods are applied. Many studies on giving users their favorite TV program information using EPG and user habits [10, 26, 27] cannot measure the popularity of TV programs. Since descriptions of television program are not always text-based, text filtering technique cannot be applied directly; clustering techniques are thus generally applied to television recommendation systems.

The present study proposes a personalized popular DVB-T digital TV program recommendation system for P2P social networks. The Multimedia Home Platform (MHP) is used to store the watching habits of users with Ratings Sharing server. The 3PRS system connects to a P2P social network; this system puts each user's interests onto a server. The server then determines related peers and automatically forms a social network.

Using Information Retrieval (IR), we assign the duration of the program that has been watched to the program weight and use a clustering method to group users. The weight of the program indicates its popularity. The clusters of users are grouped by k-means to find cluster relationships. When a new user watches a program, current clusters of users are extracted from the Ratings Sharing server. The kNN classification method is then used to determine the predicted cluster label of the present user. Popular programs in the predicted cluster and similar clusters are recommended to the user.

In our recommendation system, the IR and clustering methods are utilized to group users that have similar habits of watching TV programs into a group. IR is used to assign the weight of an item according to its attributes. A commonly used attribute is frequency: items that appear in high frequency are important, so they are assigned higher weights. However, in each user record, the frequency of a program that has watched cannot represent its popularity, so the duration of the program that has been watched is used as its weight. Thus, user records can be converted to vectors. K-means is used to cluster the users, who are represented as vectors. Popular programs in each cluster are defined based on their watched duration. When a new user appears, his or her watching habits are converted into a vector. The predicted cluster label is determined using the kNN classification method. kNN searches for k vectors that are the closest to the unlabeled user, and assigns the cluster label with the largest number in k to the user, which is called a max-win mechanism. After determining the users' predicted cluster, the system can recommend popular programs in the predicted cluster to the user.

The primary scientific contributions of the paper are as follows:

1. When using EPG, users can easily obtain information of available program such as program names, channel numbers, and their broadcast time. etc. When users select programs though EPG, we can disorderly record the length of time which users watching a program, then we analyze the recorded information to provide some input

- for recommendation systems. In this case, a recommendation system can provide an index of popularity to other users.
2. The IR technique is used to build up the recommendation system to analyze the user' interest cluster; the IR technique can be used to recommend popular programs in each cluster.

The rest of this paper is organized as follows. Related technologies are reviewed in Section 2. In Section 3, we describe the full architecture of 3PRS and the system architecture of Digital Television Client and Rating Sharing Server. Section 4 describes the implementation and shows the results. We summarize the contributions of this paper in Section 5.

2 Related work

2.1 Electronic program guide

EPG is a kind of television program list which is sent along with the digital television signal or digital broadcast signal. These signals can be received by cable TV, satellite TV, or ground TV. The most popular approach to browsing EPG [4, 11] is using a television and Set Top Box. EPGs usually have a graphical user interface which contains information like program name, channel, and play time. Users can view detailed information about any program. Broadcasting programs may contain a simple text-based message. For example, users can search for programs according to program name and channel, locate the selected program quickly, use a program reminder, and set parental controls. It also supports Void on Demand (VOD) [15], using the authorization or PPV method. EPG can also be used to perform program recording. EPG is usually broadcast along with the same channel which the TV program information being transferred, or through a dedicated message channel. For example, the ATSC standard in digital TV uses Program and System Information Protocol (PSIP) to send EPG. If only the simulated signal is available, some systems, such as TiVo, will use a modem to receive the EPG message from a third party. The device receives time signals from the local broadcast service provider so that it can record programs at some particular time. As shown in Fig. 1, broadcasting programs may contain a simple text-based message.

2.1.1 K-means

K-means [3] is an unsupervised learning [1, 3] method. It identifies the structure of data distribution from given data. Machine learning has been widely applied to many information problems, such as market segmentation [16], recommendation systems [14], image segmentation [20], privacy-preservation [9], and dimension reduction [25].

Given an integer k , n points, and $X = \{x_1, x_2, \dots, x_n\}$, k -means clusters n points into k groups as follows:

1. Generate k centers randomly, which are $C = \{c_1, c_2, \dots, c_n\}$.
2. Assign $x_i \in X$ to the nearest center $x_j \in C$
3. Calculate the object value: $\sum_{i=1}^n \min_{1 \leq j \leq k} (x_i - c_j)^2$. If there is no convergence, each center is regenerated based on the points that belong to it, then repeat step b and c