IPTV-VOD Program Recommendation System 
Using Single-Scaled Hybrid Filtering

KYUSIK PARK, JONGMU CHOI
Department of Computer Science
Dankook University
125 Jukjeon-dong, Suji-gu, Yongin-si, Gyeonggi-do
KOREA
kspark@dankook.ac.kr, choijm@dankook.ac.kr

DONGHEE LEE
School of Computer Science
Univ. of Seoul
13 Siripdae-gil, Dongdaemun-gu, Seoul
KOREA
dhl.exp.mail@gmail.com

Abstract: - In this paper, a new program recommendation algorithm is proposed to recommend user preferred VOD program in IPTV environment. A proposed system is implemented with hybrid filtering method that can cooperatively complements the shortcomings of the content-based filtering and collaborative filtering. For a user program preference, a single-scaled measure is designed so that the recommendation performance between content-based filtering and collaborative filtering is easily compared and reflected to final hybrid filtering procedure. In order to provide a high quality of program recommendation, we use not only the user watching history, but also the user program preference and mid-subgenre program preference updated weekly as a user preference profile. System performance is evaluated with modified IPTV data from real 24-weeks cable TV watching data provided by Nilson Research Corp. in Korea and it shows quite comparative quality of recommendation.

Key-Words: - IPTV, Recommendation system, Hybrid filtering, Content-based filtering, Collaborative filtering, program preference

1 Introduction

IPTV (Internet Protocol TV) is a next-generation media that can provide convergence service of broadcasting and communication using high speed internet. The main role of IPTV is to offer personalized program contents via VOD (Video on Demand) at anytime, anywhere. For this reason, a study of personalized program recommendation algorithm in IPTV environment has been broadly focused recently.

As a recommendation algorithm, there are two major approaches for general recommendation systems, i.e., the content-based filtering (CBF) approach and the collaborative filtering (CB) approach. In the content-based filtering, the user profiles are first formed by extracting features of items or programs which have been accessed in the past. Based on the user profiles, the system recommends only the items that are highly relevant to the user profiles by computing the similarities between items and the user profiles. Examples of such systems are NewsWeeder [1], Infofinders [2]. In this approach, the feature accuracy for representation of item contents is key issue which dominates the effectiveness of the recommendation. However, content-based recommendation system only able to recommend the items or programs in which a user has indicated his/her interest in past. On the other hand, the collaborative approach computes the similarities between the user profiles. Users of similar profiles are grouped together to share the information in there profiles. The main goal of the collaborative approach is to make the recommendation among the users in the same group. Examples of such system are Ringo [3] and Siteseet [4].

Adopting the collaborative filtering approach, the system has a high possibility to recommend surprising items or programs by the nature of information sharing which cannot be achieved by the content-based filtering approach. But, it has well-known cold start problem.

In this paper, we propose a new IPTV-VOD program recommendation system based on hybrid filtering algorithm. In comparison with previous works, the major contributions of this paper can be summarized as follows.
As a recommendation algorithm, a hybrid filtering that can cooperatively complements the shortcomings of the content-based filtering and collaborative filtering is implemented. For this purpose, a single-scaled measure that can compute user preference is designed to compare recommendation performance between content-based filtering and collaborative filtering, and feed this result into final hybrid filtering procedure.

Not like the most previous hybrid filtering works [5][6][7], that recommends a fixed rate of programs from content-based filtering and collaborative filtering, the proposed system does not care whether the recommendation comes from content-based filtering or collaborative filtering. It just simply compares the user preference scores using single-scaled measure to provide effective program recommendations.

As a user preference profile which is a main key for the success of recommendation algorithm, a user profile is adaptively constructed to reflect the most current user interest on a weekly basis and it gives a weight to consecutive watching for series program. It leads a better personalized recommendation.

Since IPTV watching data has not been released to open public so far, real-24weeks cable TV watching data provided by Nilson Research Corp. in Korea is modified and reconstructed to fit IPTV environment and the proposed system is evaluated with this modified IPTV data.

This paper is organized as following. Section 2 describes overall proposed IPTV recommendation system. A single-scaled hybrid filtering algorithm is fully investigated in section 3. Experimental environments and corresponding system performance are analyzed in section 4. Finally, the paper ends with conclusion in section 5.

2 Structure of the Proposed IPTV Program Recommendation System

Fig. 1 shows the main architecture for the personalized IPTV-VOD program recommendation agent in proposed system.

As seen in Fig. 1, a program agent is mainly composed of Set-Top Box (STB) and the recommendation server. Set-Top Box in fig.1 plays a role of receiving IPTV-VOD program list from the server, providing a list of program recommendations to user through UI (User Interface), and also transmit user watching history to the server. On the other hand, the server consists of user model management engine, recommendation engine, watching history DB, user model DB, and broadcasting program DB. Watching history DB stores individual user watching history received from a Set-Top Box. User model management engine computes mid-subgenre preference and the program preference on a weekly basis and store these results into user model DB. On the other hand, broadcasting program DB stores a current IPTV-VOD programs being telecast which is also updated weekly. As a final recommendation, a hybrid filtering method based on user model DB, mid-subgenre preference and program preference is applied to IPTV-VOD program DB to recommend top-five programs ordered in final program preference.

2.1 Broadcasting program DB

Broadcasting program DB stores current IPTV-VOD program as a metadata which is composed of program ID, program name, broadcasting time, program genre, and mid-subgenres. In this paper, we categorized IPTV-VOD program into 13 genres and 60 mid-subgenres.

2.2 Watching history DB and User model DB

A list of received IPTV-VOD programs from the server and the actual user watching information from the recommendation list are reorganized as a final watching history to be stored in watching history DB. On the other hand, a user model DB stores program preference and mid-subgenre preference values which is weekly updated in user model management engine.

2.3 User model management engine

User model is a main key to recognize a user watching pattern of the IPTV-VOD programs and the model computes program preference and mid-subgenre preference based on watching history information as in fig.2. A user program preference is calculated using only the most recent two-weeks watching history information. In order to reflect the most current user watching history
and to give a weight to consecutive watching for series program, a user program preference and mid-subgenre preference is updated weekly also.

### 2.3.1 Computation and update of program preference

Program preference for user \( u_i \) can be calculated by referring a recent two-weeks watching history. It is defined as a normalization of watching frequency \( p_j \) regarding program \( j \), \( j=1,2,3,...,N \). where \( N \) is total number of program, by maximum frequency \( \max_j(p_j) \). In order to give a weight for watching of series program, we give a weight of \( w_j=1\) when a user watch a series program in two consecutive weeks, otherwise, it gets \( w_j=1/2 \). Equation (1) shows program preference score (PPS) \( PPS_{j}^{ui} \) for user \( u_i \) to program \( j \) and it has values between 0 and 1.

\[
PPS_{j}^{ui} = \frac{p_j w_j}{\max_j(p_j)} , \quad j=1,2,3,...,N \quad (1)
\]

In order to reflect current user interest, a user program preference is updated on a weekly basis by considering previous and current program preference simultaneously. In other words, updated program preference for user \( u_i \) to program \( j \) in \( m^{th} \) week is given as average of program preference of \( m^{th} \) week and \( (m-1)^{th} \) week as in Eq. (2).

\[
PPS_{j}^{ui}_{\text{update}} = \frac{PPS_{j}^{ui}(\text{week}_{m-1}) + PPS_{j}^{ui}(\text{week}_{m})}{2} \quad (2)
\]

In equation (2), \( PPS_{j}^{ui}(\text{week}_{m-1}) \), \( PPS_{j}^{ui}(\text{week}_{m}) \) are program preferences of \( m^{th} \) week and \( (m-1)^{th} \) week for user \( u_i \) to program \( j \) and \( PPS_{j}^{ui}_{\text{update}} \) is updated program preference for \( m^{th} \) week that reflects program preference of previous week.

### 2.3.2 Computation and update of mid-subgenre preference

Mid-subgenre preference indicates a user interest to IPTV-VOD program subgenre and it is calculated and updated in every week. Mid-subgenre preference for a user \( u_i \) is defined as a normalization of watching frequency \( g_k \) regarding program subgenre \( k \), \( k=1,2,3,...,L \). where \( L \) is a total number of program subgenre, by maximum frequency \( \max_k(g_k) \). Equation (3) shows class preference score (CPS) \( CPS_{k}^{ui} \) for user \( u_i \) to program subgenre \( k \) and it has values between 0 and 1.

\[
CPS_{k}^{ui} = \frac{g_k}{\max_k(g_k)} , \quad k=1,2,3,...,L \quad (3)
\]

Mid-subgenre preference is updated every week by taking an average of mid-subgenre preference of \( m^{th} \) week and \( (m-1)^{th} \) week as in Eq. (4).

\[
CPS_{k}^{ui}_{\text{update}} = \frac{CPS_{k}^{ui}(\text{week}_{m-1}) + CPS_{k}^{ui}(\text{week}_{m})}{2} \quad (4)
\]

### 3 Proposed Recommendation Algorithms for IPTV-VOD Program

The proposed program recommendation algorithm uses a hybrid filtering method that can cooperatively complement the content-based filtering and collaborative filtering. At first, final program preference in each recommendation filtering is calculated as a multiplication of a program preference and mid-subgenre preference as described in section 2.3. Each filtering method then recommends five programs according to the final program preference score. Finally, the proposed hybrid filtering compares the preference score of recommended programs from each filtering and pick out top-ranked five programs.

### 3.1 Calculation of final program preference

#### 3.1.1. Content-based filtering (CBF)

Content-based filtering recommends the programs in which the user has indicated his/her interest in previous watching history. In this paper, content-based filtering calculates final program preference as a multiplication of program preference and mid-subgenre preference, and it recommends top-five ranked programs ordered in final preference score. Final preference for user \( u_i \) to program \( j \) in \( m^{th} \) week can be calculated as in Eq. (5) and it is updated every week.

\[
CBF_{j}^{ui}(\text{week}_{m}) = \frac{PPS_{j}^{ui}_{\text{update}}(\text{week}_{m}) \times CPS_{k}^{ui}_{\text{update}}(\text{week}_{m})}{PPS_{j}^{ui}(\text{week}_{m-1})} \quad (5)
\]

In equation (5), \( CBF_{j}^{ui}(\text{week}_{m}) \) is a final program preference for program \( j \) in \( m^{th} \) week calculated by content-based filtering. \( PPS_{j}^{ui}_{\text{update}}(\text{week}_{m}) \) is updated program preference for program \( j \) in \( m^{th} \) week, and \( CPS_{k}^{ui}_{\text{update}}(\text{week}_{m}) \) is updated mid-subgenre preference for program \( j \) in mid-subgenre \( k \).

#### 3.1.2. Collaborative filtering (CF)

Collaborative filtering recommends programs by referring the other users with similar interest. Program recommendation by collaborative filtering consists of two steps. First step is to group the neighbor users who
have similar watching patterns. Then the algorithm calculates final program preference by taking average program preference of neighbor users calculated in same ways of equation (5). Final program recommendation is to pick up top-five ranked programs.

Equation (6) calculates user similarity (US) to group the neighbor users with similar watching patterns and it forms five users in one group by lining up in order of US.

\[
US(u_i, u_j) = 1 - \sum_{l=1}^{G} (|GR_{ui}^l - GR_{uj}^l|), \ i \neq j \tag{6}
\]

Here, \( u_i \) is a user who will be recommended, \( u_j \) is neighbor user, \( G \) is a total number of mid-subgenres which is 60 in this paper. \( GR_{ui}^l \) is a mid-subgenre preference ratio for user \( u_i \), so the user similarity (US) will be close to 1 as the mid-subgenre preference of a user \( u_i \) and neighbor user \( u_j \) is similar. User group is also updated weekly.

Final program preference by collaborative filtering is calculated as an averages preferences in remaining four users in a group. Here the preference of each user is also calculated as a multiplication of program preference and mid-subgenre preference as in Eq. (5). Eq. (7) shows a final program preference for user \( u_i \) to program \( j \) which is updated in every week.

\[
CF_{ji}^u(week_m) = \frac{\sum_{n=0}^{40} CPS_{update_n j}(week_m) \times US_{week_m}(u_i, u_j)}{4} \tag{7}
\]

In Eq. (7), \( CF_{ji}^u(week_m) \) is a final program preference for user \( u_i \) to program \( j \) in \( m \)th week by collaborative filtering, and \( CPS_{update_n j}(week_m) \) is updated program preference score of program \( j \) for remaining four users excluding a user \( u_i \). And \( CPS_{update_k j}(week_m) \) is an updated mid-subgenre preference of program \( j \) to mid-subgenre \( k \) in \( m \)th week. \( US_{week_m}(u_i, u_j) \) is a user similarity between user \( u_i \) and \( u_j \) in \( m \)th week.

3.2 Recommendation engine of the proposed hybrid filtering method

IPTV-VOD recommendation engine in this paper utilizes a user program preference and mid-subgenre preference to recommend IPTV programs. This engine cooperatively complements the content-based filtering and collaborative filtering, so it can not only recommend various kinds of IPTV programs, but also possible to recommend for the new IPTV program.

Content-based filtering calculates final user preference as a multiplication of program preference and mid-subgenre preference and it recommends top-five ranked programs. On the other hand, collaborative filtering first group neighbor users who have similar watching patterns and then choose top-five ranked programs by taking average of final program preferences of remaining four users in a group. As seen in Eq. (5) and (7), the final program preference for both content-based filtering and collaborative filtering is scaled between 0 and 1, and so, we can compare recommendation performance between these two filtering methods using a single-scaled measure and reflect this result into final hybrid filtering procedure. In other words, the proposed hybrid filtering algorithm compares ten program preferences submitted from each filtering method, then it ultimately recommends top-five ranked IPTV programs to the users.

Figure 2 shows the flowchart of the proposed hybrid recommendation algorithm.

4 Experimental Environment and Simulation Results

4.1 Experimental environment
Since IPTV watching data is not opened to public yet, we used cable TV watching data provided by Nilson Research Corp. in Korea. This data was modified and reconstructed to fit IPTV environment to evaluate the system performance. Nilson Research data used in this paper is real 24-weeks watching data from Jan. 1st of 2008 to June 14th of 2008 for terrestrial broadcast and cable TV.

After modification of the Nilson Research data, the first four-week data is used as a training data, and the rest of data is used as a test data for the experiment. Training data is used to calculate the initial program preference and mid-subgenre preference in content-based filtering and also it is used to form initial neighbor group in collaborative filtering. Test data is used to evaluate the system performance.

Followings are major modifications made to Nilson Research data to fit for IPTV-VOD environment. Firstly, to match the characteristics of the IPTV-VOD programs, one-time broadcasting programs such as news, weather, traffic, etc. was removed from the Nilson Research data and only the rest of programs were used to build IPTV-VOD program DB. Secondly, the program genre classification was made by following the program classification system made by KT, one of the well-known IPTV service providers in Korea and as a result, we have a total of 13 program genre with 60 mid-subgenres. Finally, to get accurate statistics of user program preference and mid-subgenre preference, we only consider the programs that a user watches more than 10% of the program to build user model DB.

4.2 Experimental results and analysis

There are many ways to evaluate the recommendation system and such are Hit ratio, Ranking point average, Percentage of contents, Precision, and Recall, etc. In this paper, we used Precision as a performance evaluation measure and it describes percentage of the actual user watching based on five recommended programs made by the proposed IPTV-VOD recommendation engine. Eq. (8) shows the mathematical formula for the Precision measure. Here, N(=5) is a total number of recommended programs by the system and N_k is the number of programs a user actually watched from the recommendation list.

\[
\text{Precision} = \frac{N_k}{N}
\]  

(8)

The proposed system was evaluated in followings three aspects. Firstly, as in the most previous research works, the hybrid system with a fixed ratio of recommendations from the content-based filtering and collaborative filtering was compared. Secondly, the proposed hybrid recommendation system was evaluated and compared in terms of Precision measure. Finally, the proposed system was analyzed with different number of program recommendations.

4.2.1 Performance comparison of hybrid system with a fixed ratio of recommendations from the content-based filtering and collaborative filtering

Figure 3 compares the system performance when the number of recommendations from the content-based filtering and collaborative filtering has various fixed ratios 4:1, 3:2, 2:3, and 1:4. For example, 4:1 fixed ratio of recommendation means that the system take four recommendations from the content-based filtering and take only one recommendation from the collaborative filtering.

Fig. 3 Performance comparison of fixed ratio hybrid system

As seen on the figure, the recommendation performance is best when the recommendation ratio is 4:1. It means that the content-based filtering has more influence to overall system performance. On the other hand, we notice that the hybrid system in figure 5 shows a little bit of improved performance than the one with individual filtering method in section 4.2.1. It actually explains the efficiency of the hybrid system.

4.2.2 Performance comparison of the proposed hybrid recommendation system

Not like the previous hybrid system in section 4.2.2 which recommends a fixed ratio of programs from content-based filtering and collaborative filtering, the proposed system does not care whether the recommendation comes from content-based filtering or collaborative filtering. It just simply compares the final program preference score using a single-scaled measure to provide top-five ranked program recommendations.

Figure 4 compares system performance of the collaborative filtering in section 4.2.1, hybrid system with fixed ratio of 4:1 in section 4.2.2 and the proposed hybrid filtering system.
Fig. 4 Performance comparison of the proposed hybrid recommendation system

From the figure, the proposed system outperforms all other methods with a average precision of 0.72 which is 4.3% improvement over the collaborative filtering in section 4.2.1 and 1.7% improvement over the system in section 4.2.2. Such a performance improvement comes from the flexible program recommendation ability with a single-scaled program preference scale.

4.2.3 System performance with different number of program recommendations

Figure 5 compares system performance when we vary the number of program recommendations to 5, 10 and 15. From the figure, we see that the system performance is gradually dropped from 0.72 to 0.556 as the number of program recommendation is increased from 5 to 15. For this reason, all the experiments in this paper were performed with an assumption of five recommendations.

Fig. 5 System performance with different number of program recommendations

5. Conclusion

This paper proposes a new IPTV-VOD program recommendation algorithm with flexible hybrid filtering method. A proposed system works complementally to cooperate the content-based filtering and collaborative filtering. A single-scaled measure is designed to calculate final user program preference and it is used to compare the program preference of each filtering method. Then this comparison result is reflected to final hybrid filtering procedure for IPTV-VOD program recommendation. In order to provide high quality of program recommendation, not only the user watching history, but also the user program preference and mid-subgenre program preference are updated in every week and they are reflected to a final user program preference profile. System performance is evaluated with modified IPTV-VOD data provided by Nilson Research Corp. in Korea and it shows quite comparative quality of recommendation compared to previous research results. The proposed system can provide very flexible and stable program recommendations. For the future research, a further elaboration of the proposed system will be investigated so that the system can adaptively adjusted to a given system conditions.

References: