

사용자 선호도 기반의 TV 프로그램 추천 기법

윤현호⁰, 이성진, 강영길, 이수원, 김현*

송실대학교 대학원 컴퓨터학과

{likechunhee⁰, ptnrev93, dudrlf34, swlee}@mining.ssu.ac.kr

User Preference based personalized Electronic Program Guide

Hyunho Yoon⁰, Younggil Kang, Seongiin Lee, Soowon Lee

Dept. of Computing, Graduate, School, Soongsil University

요약

디지털 방송이 본격적으로 시작되고 각 가정에 디지털 TV의 보급이 많이 이루어짐에 따라 사용자들이 시청할 수 있는 채널 및 프로그램의 수는 기존의 공중파 환경과 비교하여 폭발적으로 증가하였다. 채널의 수가 수백 개로 늘어나기 때문에 사용자가 자신이 원하는 프로그램을 찾기 위해서는 리모콘을 누르거나 방송 편성표를 검색하는 것과 같은 많은 시간과 노력을 필요하다. 따라서 이러한 문제점을 해결하기 위해서는 사용자의 프로파일과 선호도 정보를 분석하여 사용자의 취향에 적합한 프로그램을 추천해줄 수 있는 방법이 필요하다. 본 논문에서는 수백 개의 채널에서 방송되는 프로그램 및 사용자의 선호도 정보를 분석하여 사용자가 원하는 프로그램을 추천해주고 시청 히스토리를 분석하여 사용자 모델을 동적으로 갱신하는 기법을 제안한다.

1. 서론

토리를 분석하여 사용자 모델을 동적으로 갱신하는 기법을 제안한다.

디지털 방송이 본격적으로 시작되고 각 가정에 디지털 TV의 보급이 많이 이루어짐에 따라 사용자들이 시청할 수 있는 채널 및 프로그램의 수는 기존의 공중파 환경과 비교하여 폭발적으로 증가하였다. 기존의 공중파나 케이블 TV에서 제공하는 프로그램의 수는 그리 많지 않기 때문에 사용자가 자신이 원하는 프로그램을 실시간으로 찾아서 시청하는 것은 그리 큰 문제가 아니었다.

2. 관련 연구

2.1. 일반적인 추천 기법

현재 많이 사용되고 있는 대표적인 추천 기법은 내용 기반 추천(Content-based Recommendation)과 협업 추천(Collaborative Recommendation)이다. 내용 기반 추천은 정보 검색(Information Retrieval) 분야에 바탕을 두고 있으며 사용자의 프로파일과 추천의 대상이 되는 아이템의 구성 요소를 비교하여 유사도가 높은 것들을 추천하는 기법이다. 협업 추천은 추천의 대상이 되는 목표 사용자와 비슷한 프로파일을 가진 다른 사용자를 찾아 그 사용자가 높은 평가를 매긴 아이템을 목표 사용자에게 추천한다[1].

그러나 디지털 TV 환경 하에서는 채널의 수가 수백 개로 늘어나기 때문에 사용자가 자신이 원하는 프로그램을 찾기 위해 리모콘을 누르거나 방송 편성표를 검색하는 것은 많은 시간과 노력을 필요로 한다. 즉 어떠한 프로그램을 볼 것인가에 대한 선택의 문제가 대두되게 된다.

이 외에 사용자의 성별, 나이, 직업 등의 인구통계학적 정보를 활용하여 추천을 하는 인구통계학적 추천 기법과 내용 기반 추천과 협업 추천을 결합한 하이브리드 추천 기법 등이 있다.

따라서 이러한 문제점을 해결하기 위해서는 사용자의 프로파일과 선호도 정보를 분석하여 사용자의 취향에 적합한 프로그램을 추천해줄 수 있는 방법이 필요하며, 이러한 요구의 일환으로 개인화된 방송 편성표(pEPG : personalized Electronic Program Guide)에 대한 연구가 활발히 진행되고 있다. pEPG는 필터링 기법을 활용하여 사용자의 취향에 맞는 프로그램을 제공하는 시스템을 말한다. 본 논문에서는 수백 개의 채널에서 방송되는 프로그램 및 사용자의 선호도 정보를 분석하여 사용자가 원하는 프로그램을 추천해주고 시청 히스

2.2. 기존의 pEPG에 관한 연구

현재 제품화되어 있는 pEPG로는 tivo가 있다. tivo(<http://www.tivo.com>)는 PVR(Personal Video Recorder)을 활용하여 사용자가 선호하는 프로그램을 자동으로 찾아서 저장한다. 그러나 사용자의 명시적인 선호도 정보와 프로그램 정보의 단순 비교를 통한 추천 결과를 제시한다.

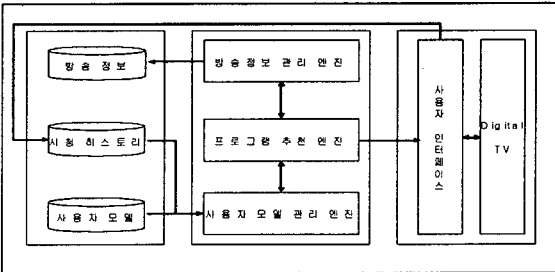
* 본 연구는 한국전자통신연구원의 지원으로 이루어졌습니다.

[2][3]은 내용 기반의 추천을, [4]는 인구통계학적 추천 기법과 협업 추천 및 내용 기반 추천을 혼합한 추천 방식을 제안하고 있다. [5]는 웹기반의 pEPG로서 하이브리드 추천 방식을 제공하며, [6]은 사용자의 시청 습관에 따라 미리 정의한 스테레오타입(Stereotype)으로 분류하고 각 스테레오타입에 맞는 추천 리스트를 제공한다. [7] 역시 베이지안(Baysian) 알고리즘을 활용한 내용기반의 추천 기법을 사용하며 사용자의 명시적, 묵시적 피드백을 추천에 반영한다.

본 논문은 여러 추천 기법 중 내용 기반의 추천 기법을 바탕으로 하여 사용자의 요일별/시간대별 시청 장르에 대한 빈도를 분석한 후 사용자의 취향에 맞는 TV 프로그램을 추천해 준다. 동시에 시청 히스토리를 분석하여 사용자 모델을 동적으로 갱신하는 기법을 제안한다.

3. 추천 시스템

본 논문에서 제안하고 있는 추천 시스템의 구조도는 [그림 1]과 같다.



[그림 1] 방송 프로그램 추천 시스템의 구조도

본 시스템에서는 사용자의 선호도 정보와 방송 프로그램 정보간의 유사도를 계산하여 사용자가 흥미를 가질만한 방송 프로그램의 목록을 제공한다. 방송 프로그램 정보와 사용자의 선호도 정보를 저장하기 위해 방송 정보 저장소와 사용자 모델을 필요로 하며 사용자의 선호도를 동적으로 갱신하기 위해 사용자의 시청 정보를 시청 히스토리(watching history) 저장소에 저장한다.

3.1 사용자 모델

사용자 모델은 TV 프로그램에 관련된 장르, 채널, 인물 등에 관한 선호도를 반영하는 TV Specific Model[표 1]과 전반적인 영역에서의 사용자의 관심사(키워드)를 반영하는 공통 모델(Common Model)[표 2]로 구성된다.

또한 사용자마다 요일별/시간대별 장르에 따라 서로 다른 시청 패턴을 보이며 TV 프로그램의 경우 평일과 주말의 편성 특징이 크게 다르므로 이를 반영하기 위해 요일별/시간대별 시청 장르에 대한 빈도 매트릭스[표 3]를 TV Specific

Model에 저장한다.

[표 1] TV Specific Model의 구성

선호 세부 장르	<세부장르1, 선호도>, <세부장르2, 선호도>
선호 채널	<채널1, 선호도>, <채널2, 선호도>
인물 키워드 벡터	<인물1, 선호도>, <인물2, 선호도>

[표 2] 공통 모델의 구성

<키워드1, 선호도>, <키워드1, 선호도>, ...

[표 3] 요일별/시간대별 시청 장르에 대한 빈도 매트릭스

	0시 ~ 1시				...	23시 ~ 0시			
	장르1	장르2	...	장르N	...	장르1	장르2	...	장르N
평일	Freq.	Freq.	Freq.	Freq.	...	Freq.	Freq.	Freq.	Freq.
토요일	Freq.	Freq.	Freq.	Freq.	...	Freq.	Freq.	Freq.	Freq.
일요일	Freq.	Freq.	Freq.	Freq.	...	Freq.	Freq.	Freq.	Freq.

3.2 사용자 모델에 의한 프로그램의 선호도 계산

사용자별 프로그램에 대한 선호도의 계산은 TV Specific Model과 공통 모델의 선호도를 계산하여 합한 값으로 이루어진다.

$$Preference^{u,p} = \alpha \times Preference^{u,p}_{Specific} + \beta \times Preference^{u,p}_{Common}$$

- u : 사용자, p : 방송 프로그램
- Preference^{u,p}: 사용자 모델에 의한 사용자 u의 p에 대한 선호도
- Preference^{u,p}_{Specific}: TV Specific Model에 의한 선호도
- Preference^{u,p}_{Common}: 공통 모델에 의한 선호도
- α : TV Specific Model에 대한 가중치
- β : 공통 모델에 대한 가중치

[식 1] 사용자 모델의 의한 프로그램의 선호도 계산

TV Specific Model에 의한 선호도는 장르, 채널, 인물에 의한 선호도를 계산한 후 각각의 가중치를 곱한 값의 합으로 계산된다.

$$Preference^{u,p}_{Specific} = w_G \times Preference^{u,p}_{Genre} + w_C \times Preference^{u,p}_{Channel} + w_P \times Preference^{u,p}_{Person}$$

[식 2] TV Specific Model에 의한 선호도의 계산

공통 모델에 의한 선호도 계산은 방송 프로그램의 제목, 내용, 개요, 줄거리로부터 추출한 키워드 벡터와 사용자 키워드 벡터와의 코사인 유사도를 구한다.

$$Preference_{Person}^{u,p} = COSINE_SIM(Vector_{PersonOpp}, Vector_{PersonOfModel})$$

[식 3] 공통 모델에 의한 선호도 계산

3.3 사용자 모델의 갱신

요일별/시간대별 시청 장르에 대한 빈도 매트릭스는 사용자의 시청 히스토리에서 요일별/시간대별로 시청한 요일, 시간, 장르를 분석하여 해당 매트릭스 횡수를 증가시킨다.

TV Specific Model의 갱신은 먼저 일정 기간 이상 시청하지 않는 세부 장르나 채널, 인물을 삭제하기 위하여 각각의 선호도를 0.1씩 감소시킨다. 다음으로 사용자의 시청 히스토리로부터 긍정 피드백과 부정 피드백을 구분하여 세부 장르, 채널, 인물의 출현 빈도에 각각의 가중치를 곱한 값을 계산한 후 각각의 Max 값을 이용하여 정규화하여 갱신할 값을 계산한다. 마지막으로 모델에 저장되어 있지 않은 내용은 새로이 추가하고 이미 존재하는 내용에 대해서는 기존의 선호도에 갱신 값을 더한다. 갱신 후의 선호도가 0일 때에는 모델에서 삭제한다.

$$V_{sg}^u = \sum_{i \in F(+)} P_{subGenre} + \sum_{i \in F(-)} N_{subGenre}, sg \in subGenre$$

$$V_{sg}^u = \frac{V_{sg}^u}{\max \{ V_{sg}^u, sg \in subGenre \}}$$

$$Preference_{sg}^u = Preference_{sg}^u + V_{sg}^u, sg \in subGenre$$

u : 사용자

F(+): 사용자가 positive로 피드백을 준 프로그램의 집합

F(-): 사용자가 negative로 피드백을 준 프로그램의 집합

subGenre : 프로그램의 세부 장르

P_{subGenre} : positive 프로그램의 세부 장르에 대한 가중치

N_{subGenre} : negative 프로그램의 세부 장르에 대한 가중치

[식 4] 세부 장르에 대한 선호도 갱신

채널과 인물에 대한 갱신 방법은 위와 동일하다.

공통 모델의 선호 키워드를 갱신하기 위해서는 우선 가장 최근의 관심사를 반영하기 위하여 공통 모델에 저장되어 있는 모든 키워드의 선호도를 0.1씩 감소시킨다. 이하 공통 모델의 갱신 방법은 [식 4]와 동일하다.

4. 실험 및 평가

본 논문에서는 실험을 위해 2005년 3월부터 30일간 지상파, 케이블, 위성 방송을 포함한 133개 채널로부터 방송되는 모든 채널을 대상으로 하였으며 총 3명이 실험에 참가하였다. 최대 추천 목록은 10개, 업데이트 주기는 1일로 하였으며, α : 0.8, β : 0.2, w_G : 0.56, w_C : 0.14, w_P : 0.3으로 설정하였다.

사용자의 시청 히스토리 기록은 시청함과 시청하지 않음의 두 개의 값으로 기록하였으며, 추천 프로그램에 대한 피드백 역시 긍정과 부정의 두 가지 척도로만 입력 받았다. 초기 사용자 모델을 모두 동일하게 설정하여 시청 히스토리에 따라 사용자의 선호도 정보와 추천 결과가 어떻게 변화하는지를 관찰하였다.

실험 결과 초기 선호도 정보가 동일하더라도 주로 시청하는 프로그램의 성격이 다를 경우 사용자의 선호도 정보가 동적으로 변화하고 추천 결과가 서로 달라짐을 확인할 수 있었다.

5. 결론

본 논문은 디지털 TV 환경 하에서의 개인화된 방송 편성표 (pEPG)에 관한 연구로서 사용자의 선호도 정보를 분석하여 사용자의 취향에 적합한 방송 프로그램을 제공하고 시청 히스토리를 반영하여 동적으로 사용자 모델을 갱신하고 이를 추천 결과에 반영하는 기법을 제안하였다.

향후에는 추천의 정확도를 보다 향상시키기 위해 개인적 취향에 의한 추천뿐 아니라 시청률을 추천에 반영하고 스테레오 타입 구축 등을 통하여 같은 그룹에 속하는 사용자들의 시청 특징을 반영한다면 보다 나은 추천 결과를 얻을 수 있을 것이다.

6. 참고 문헌

- [1] Marko Balabnovic and Yova Shoham, Fab:Content-Based, Collaborative Recommendation, CACM 40(3) p66-72, 1997
- [2] 유상원, 내용 기반 추천 기법의 TV 환경 적용에 관한 연구, 한국정보과학회학술발표논문집, 2483호, p797-799, 2003
- [3] 류지용, 사용자 선호도 기반 지능형 프로그램 가이드, 한국방송공학회 논문지, 제 7권 2호, 2002
- [4] 송서리, 개인화된 TV 서비스를 위한 추천 기법 개선, 한국정보처리학회 추계학술발표대회 논문집, 제 11권 2호, 2004
- [5] P. Baudisch, Designing an evolving Internet TV program guide, In *Human Computer Interaction Consortium Winter Workshop(HCIC '97)*, Snow Mountain Ranch, Co, 1997
- [6] L. Ardissono, User Modeling and Recommendation Techniques for personalized Electronic Program Guides, In *Personalized Digital Television*, Human-Computer Interaction Series, Vol. 6, 2004
- [7] K. Kurapati, A Multi-Agent TV Recommender, In *Proceedings of the UM 2002 workshop "Personalization in Future TV"*, 2001