

# 개인화된 TV서비스를 위한 추천기법 개선

서송리\*, 배기성\*\*, 석민수\*\*  
\*성균관대학교 전기전자컴퓨터공학과  
\*\*성균관대학교 전자전기공학과  
e-mail:elegiaco@ece.skku.ac.kr

## Improving Recommendation for Personalized TV Service

Song-Lee Suh\*, Kee-Sung Bae\*\*, Min-Su Suk\*\*  
\*Dept of Electrical Computer Engineering, SungKyunKwan University  
\*\*Dept of Electronics Engineering, SungKyunKwan University

### 요 약

2001년 하반기 이후 디지털 TV 시대가 열리면서 채널의 수와 그에 따른 프로그램의 수가 폭발적으로 증가했다. 그리하여 기존의 방법으로는 시청자가 원하는 프로그램을 선택하는 것이 어려운 일이 되었다. 이 문제를 해결하는 방안으로서 pEPG(personalized Electronic Program Guide)가 많이 연구되어 왔으며 본 논문에서는 pEPG를 위한 추천 방법에 대해 연구하고자 한다. 기존의 추천 방법은 내용기반추천과 협업추천이 대표적이는데, 이들은 어느 한쪽이 우월하다기 보다 각각의 단점을 상호보완해주는 관계에 있다. 각 추천 방법이 TV환경의 pEPG에 적용될 때는 어떤 장단점이 있는지 살펴보고, 이에 인구통계학적추천을 혼합한 기법을 제안한다.

### 1. 서론

2001년 하반기 이후 디지털 TV 시대가 열리면서 흑백에서 컬러화면으로 전환하던 시기만큼 큰 변화가 찾아왔다. 시청채널 수의 증가가 원인으로서 공중파 채널 몇 개에 불과했던 것이 케이블방송, 위성방송 등은 물론 디지털 TV의 보급으로 수십, 수백 개의 채널로 폭발적인 증가를 보였다. 그리고 PVR(Personal Video Recorder)의 보급이 이루어져 지나간 방송도 볼 수 있게 됨으로서 실질적인 시청 프로그램 수는 더욱더 증가했다.

기존에는 신문 등의 프로그램 가이드를 보고 선택하거나, 즐겨보는 채널을 암기하거나, 여러 채널을 돌려보아도 별반 문제될 것이 없었다. 하지만 채널이 수백 개로 늘어나면서 시청자들은 마음에 드는 프로그램을 보기 위한 많은 노력이 필요하게 되었다. 텍스트 기반의 프로그램 가이드는 프로그램 리스트를 인쇄하기 위해 하루에 30페이지 이상이 필요하므로 더 이상 그 역할을 다 할 수가 없고, 단순히

프로그램을 채널별·시간대별로 정렬하여 제시하는 EPG(Electronic Program Guide)를 하나하나 보며 선택하는 것도 역시 효과적이지 못하다. 또한 채널을 직접 돌려보는 경우, 채널이 300개라고 가정하고 그 채널이 자기가 원하는 채널인지 아닌지를 판단하는데 1초가 걸린다고 하면 모든 채널을 탐색하는데 5분이 걸린다. 그 5분 사이에 이미 탐색한 채널의 프로그램은 새로운 프로그램으로 바뀌었을 가능성이 있으며 그것을 확인하기 위해서는 또 모든 채널을 돌려보아야 하는 상황이 발생한다.

이러한 문제점을 해결하기 위해 시청자 개인의 취향 및 선호도를 반영하여 프로그램을 추천해주는 프로그램가이드가 절실히 요구되며, 이러한 TV 추천의 한 분야로서 pEPG(personalized Electronic Program Guide)가 대두되었다. 본 논문에서는 pEPG를 위한 기존의 추천 방법에 대해 소개하고 결과 성능이 평가단계에서 기존의 기법보다 유용한 추천 기법을 제안한다.

## 2. 관련 연구

상용화되어 있는 서비스로 다음과 같은 것들이 있다. Canal plus는 1993년부터 양방향 TV서비스를 위한 미들웨어를 공급해온 회사로서 유럽 전역을 대상으로 방송사업을 펼치고 있으며 추천은 아니지만 본인이 원하는 정보를 손쉽게 검색하는 기능을 제공한다. Tivo[1]는 PVR(Personal Video Recorder)이라는 저장기능을 가진 TV를 이용해 양방향 서비스를 제공하고 있으며 원하는 키워드를 입력하면 해당 키워드가 나타나는 프로그램을 자동으로 녹화해주는 기능을 제공하고 있다.

프로그램가이드와 관련된 연구로서 [2],[3]은 사용자 선호도와 히스토리에 기반한 IPG(Intelligent Program Guide)를 제공한다. [4]는 온라인상에서 프로그램 정보를 제공하고 이를 이용하여 취향이 비슷한 사용자들끼리 묶은 후 내용기반추천과 협업추천을 혼합한 방식으로 추천리스트를 제공한다. [5]는 Agent를 기반으로 사용자에게 프로그램을 추천하는 방식을 취하고 있으며 [6]은 내용기반추천과 협업추천을 연관규칙생성에 이용하여 추천리스트를 제공한다.

## 3. 추천시스템

개인화는 이용자의 프로파일에 근거하여 내용을 선정하고, 사용자에게 새롭거나 다른 이용자가 미처 생각하지 못했던 것을 채택하는 것이다. 즉, 사용자의 특성을 대상으로 하여 이와 관련한 다양한 통계 분석 기법 및 데이터 마이닝 기법을 통해 개별화할 수 있는 지식과 규칙을 찾아내는 과정으로 볼 수 있다.

이러한 개인화를 위해서 추천시스템이 필요하다. 예전과는 달리 지금의 개인화는 단순히 사업자가 사용자에게 정보나 물품을 일방적으로 푸시(Push)하는 것이 아니라 사용자에 따른 정보를 자동으로 제공하도록 하는 것이기 때문이다. 사용자가 원하는 프로그램을 쉽게 찾음으로서 시청률을 더 높일 수 있고, 사용자의 틀에 박혀있는 프로그램 시청에 선호도가 높고 새로운 프로그램을 추천함으로써 또 다른 프로그램에 대한 충성도를 높일 수 있다. 그리고 시청자에게 지나간 프로그램이라 하더라도 추천하여 녹화한 다음 지속적으로 추천할 수 있으므로 그 결과 이익은 더 높아질 수 있다.

## 4. 기존의 추천 기법

pEPG를 위한 대표적인 추천기법으로 내용기반추천(Content-based Recommendation)과 협업추천(Collaborative Filtering)을 들 수 있으며 그 외에도 인구통계학적 추천기법, 연관성규칙기반기법, 혼합기법 등이 있다.

내용기반 추천의 기본개념은 사용자 프로파일에서 높은 유사성을 가진 프로그램을 추천하는 것이다. 이 기법은 사용자 프로파일을 통해 과거의 추천 결과를 쉽게 반영할 수 있는 장점이 있으며 추천 속도가 빠르다. 하지만 사용자와 프로그램의 정확한 묘사가 불가능한 분야에서는 내용 분석이 피상적으로 이루어질 수 있으며, 추출 결과가 유용하지 않을 수 있다. 또한, 이용자의 프로파일과 비교하여 높은 점수를 얻은 아이템을 추천하므로, 이용자가 이미 평가한 아이템과 유사한 아이템만을 제공함으로써 과도하게 특수화되는 경향을 보인다.

이러한 문제점을 해결하기 위해 협업추천을 사용한다. 협업추천은 프로그램과 사용자의 유사성을 계산하는 대신, 타겟 사용자와 비슷한 취향을 지닌 이웃 사용자를 찾아 이들의 유사성을 계산하고 이웃 사용자의 선호도 리스트에서 가장 일반적인 프로그램을 추천한다. 내용기반추천과 비교하여 협업추천은 다음과 같은 장점을 가지고 있다. 첫째, 추천할 대상의 표현에 의존할 필요 없이 검색이 가능하다. 따라서 취향에 따라 다양한 선호도를 보이는 음악이나 영화 같은 복합적인 아이템에도 잘 작동한다. 둘째, 이용자가 검색하고자 하는 내용을 포함하지 않지만 이용자에게 가치 있는 아이템을 추천할 수 있다. 그러나 협업추천도 단점이 있다. 첫째, 초기 평가자(early-rater) 문제이다. 즉, 새로운 아이템이 추가되면 이웃 사용자가 충분한 정보를 입력하거나 이용하기 전에는 그 아이템을 추천할 수가 없다. 둘째, 희박성(sparsity)의 문제이다. 대부분의 영역에서 다수의 아이템에 대해 모든 사용자가 평가할 수 없기에 이용자-아이템 행렬의 평가 값들은 매우 희박한 상태가 된다.

인구통계학적 추천은 이 기법하나만으로 이루어진 연구는 없다. 사용자의 성별, 나이, 직업 등과 같은 인구통계학적 요소에 의해 사용자 유형별 특징을 분석하여 추천하는 방법으로서 간단하고 사용자로부터의 피드백 정보가 없이도 추천이 가능하지만 추천 효율이 상당히 떨어지기 때문이다.

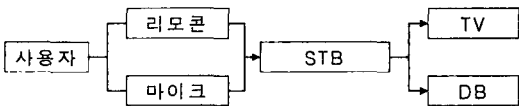
연관성 규칙 기반기법은 전자상거래에서 자주 사용되는 기법 중의 하나로 연관성 규칙을 이용하는

것이다. 이 기법은 특정 트랜잭션에 하나의 제품이 존재하고 동시에 같은 트랜잭션에 다른 제품이 존재할 때 이러한 두 제품 사이의 연관성을 발견하는 데 초점이 맞추어져 있다. 이 기법의 추천 성능은 연관성 측정 방법에 의해 큰 영향을 받으므로 오랜 기간 동안 정보를 축적하여 정확하게 연관성 정도를 산출할 수 있어야 한다.

따라서 위의 접근방법들에 대한 결점을 극복하기 위해서는 단점을 보완하고 장점을 부각시키는 방향으로 각 기법을 혼합하여 사용하는 것이 최선의 추천 방법이다.

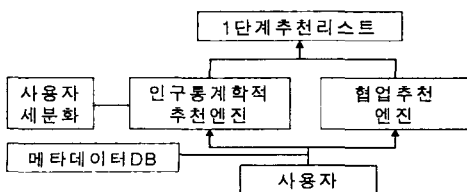
### 5. 제안기법

TV의 동적특성으로 인해 인구통계학적 추천은 이제껏 무시되어 왔으나, PVR이 상용화됨에 따라 더욱 정확한 추천을 위한 인구통계학적 추천의 사용이 가능해졌다. [그림 1]은 리모콘 뿐만 아니라 내장 마이크를 통한 음성인식 시스템을 도입하여 사용자 요구사항을 받아들이며 STB를 통한 DB의 생성과 갱신, TV 화면으로 디스플레이해주는 추천시스템의 전체적인 구성을 보여준다. 이런 시스템으로 인해 앞서 언급했듯이 인구통계학적 추천의 사용이 가능해졌으나 인구통계학적 추천만으로는 효용이 떨어지므로 이를 혼합한 3단계 방법을 제안한다.

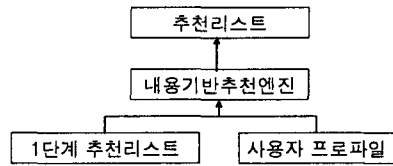


[그림 1] TV 추천 시스템

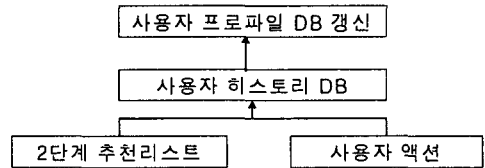
협업 추천의 초기 평가자 문제를 해결하기 위해 혼합 1단계에는 협업추천과 인구통계학적 추천을 혼합하고, 2단계로서 그 결과 리스트를 내용추천 결과와 혼합한다. 마지막 3단계에서 추천리스트에 의한 사용자 히스토리를 저장하여 프로필 DB의 갱신에 이용한다.



[그림 2] 제안기법 1단계



[그림 3] 제안기법 2단계



[그림 4] 제안기법 3단계

인구통계학적 추천에서 '나이', '성별', '직업' 등을 다른 추천기법과 하나씩 혼합하는 방법보다 '나이, 성별, 직업'을 모두 혼합하는 방법이 신뢰도가 더 높은 결과치를 보이므로[9] 혼합하는 방법과 혼합비율을 조금씩 조정하여 그 결과치를 비교해볼 수 있다.

### 6. 실험 및 분석

MPEG-7에서는 사용자 선호도 및 사용자 히스토리에 대한 기술문서를 DDL(Description Definition Language)로 표현하고 있다. TV-Anytime규격은 이러한 MPEG-7의 기술문서와 동일하게 시청자의 콘텐츠에 대한 사용자 선호도 및 사용 히스토리를 기록할 수 있는 스키마를 정의하고 있다.

TV 프로그램에 대한 정보로서 TV-Anytime[12]의 메타데이터(data about data)를 사용하여 필요부분만을 추출하여 사용할 것이며, 각각의 사용자 선호도를 기반으로 추천기법만을 사용한 결과치와 제안기법의 결과치를 비교평가하여 그 성능을 예측한다. 기법을 적용하여 나타난 결과치에 규칙기반 필터링을 사용하여 사용자 히스토리를 적용함으로써 목시적으로 충성도를 피드백하여 기법에 반영될 수 있도록 한다. 사용되는 액션 히스토리의 기술구조는 아래 [그림5][12]와 같으며 '재생, 멈춤, 빨리 감기, 되감기, 녹화' 등으로 구분하는 시청자의 행동을 모니터링하여 저장한다.

```

<!-- Definition of UserActionHistory DS (ISO/IEC 15938-5:15.3.3) -->
<complexType name="UserActionHistoryType">
  <complexContent>
    <extension base="mpeg7:DSType">
      <sequence>
    
```

```

<element name="ObservationPeriod"
  type="mpeg7:TimeType" minOccurs="0"
  maxOccurs="unbounded"/>
  <element name="UserActionList"
    type="mpeg7:UserActionListType"
    minOccurs="1" maxOccurs="unbounded"/>
</sequence>
<attribute name="protected"
  type="mpeg7:userChoiceType" use="optional"
  default="true"/>
</extension>
</complexContent>
</complexType>
<!-- Definition of UserActionList DS (ISO/IEC 15938-5:
15.3.4) -->
<complexType name="UserActionListType">
  <complexContent>
    <extension base="mpeg7:DSType">
      <sequence minOccurs="0">
        <element name="ActionType"
          type="mpeg7:TermUseType"/>
        <element name="UserAction"
          type="mpeg7:UserActionType" minOccurs="0"
          maxOccurs="unbounded"/>
      </sequence>
      <attribute name="numOfInstances"
        type="nonNegativeInteger" use="optional"/>
      <attribute name="totalDuration"
        type="mpeg7:durationType" use="optional"/>
    </extension>
  </complexContent>
</complexType>

```

[그림 5] 사용자 히스토리 기술구조

표본 집단을 선정하여 7일간 실험해볼 것이며, 현재 자동으로 히스토리를 피드백하는 것은 STB(Set-Top Box)에서만 이뤄지므로 각각은 인터넷 기반으로 프로그래밍한 추천 프로그램을 통해 수동으로 피드백하여 결과를 비교해 볼 것이다.

이와 같은 제안 기법과 기존의 기법을 비교한 결과를 제시하고, 본 연구의 기법에서도 추천의 조합 비율을 달리하여 비교한 결과를 제시할 것이다.

## 6. 결론 및 연구방향

예전에는 웹 베이스보다 생명주기가 짧은 '동적' 환경이라는 점에서 인구통계학적 추천이 불가능했다. 현재 방송중인 프로그램만을 시청 가능했기 때문이다. 하지만 PVR의 상용화가 시작되면서 방송시간에 얽매이지 않고 추천이 가능하게 되었다. 앞으로 점차적으로 PVR이 보편화될 것이므로 제안기법이 더 유용해질 것이다. 하지만 시청자마다 추천방식의 만족도가 각기 다르므로 추천의 조합비율을 적절히 변경하여 평균적인 만족도가 가장 높은 기법을

선별하여야 한다.

## 참고문헌

- [1] <http://www.tivo.com>
- [2] 배빛나라, 류지웅, 김문철, 남재호, 강경옥, 노용만, "TV Anytime 응용을 위한 사용자 선호도 추출 및 갱신 알고리즘", 한국방송공학회 학술대회, 2001년 11월 24일
- [3] 류지웅, 배빛나라, 김문철, 남재호, 강경옥, "디지털 방송을 위한 지능형 프로그램 가이드"
- [4] Barry Smyth, Paul Cotter, "Generating Personalized TV Listings Using Collaborative, Case-Based Recommendation", LNCS 1650, 1999
- [5] K. Kurapati, S. Gutta, D. Schaffer, "A multi-agent TV recommender", 2001
- [6] Derry O' Sullivan, Barry Smyth, David C. Wilson, "Improving the Quality of the Personalized Electronic Program Guide", Proceedings of the 2nd Workshop on Personalization in Future TV, May 28, 2002, Malaga, Spain.
- [7] Weizhen Dai, Robin Cohen, "Dynamic Personalized TV Recommendation System", Proceedings of the 3rd Workshop on Personalization in Future TV, June 23, 2003, Johnstown.
- [8] Jiangshan Xu, Liang-Jie Zhang, Haiming Lu, Yanda Li, "The Development and Prospect of Personalized TV Program Recommendation Systems", Proceedings of the IEEE Fourth International Symposium on Multimedia Software Engineering, 2002
- [9] 황성희, 김영지, 이미희, 우용태, "인구통계학적 특성에 따른 협동적필터링 알고리즘의 추천 효율 분석"
- [10] Philip S. Yu, "Data Mining and Personalization Technologies"
- [11] John Zimmerman, Kaushal Kurapati, Anna L. Buczak, "TV Personalization System"
- [12] <http://www.tv-anvtime.org>
- [13] Workshop on Personalization in Future TV 2001~2004, <http://www.di.unito.it/~liliana>