

TV프로그램을 위한 내용기반 추천 시스템

(A Content-based TV Program Recommender)

유 상 원 [†] 이 흥 래 [†] 이 형 동 [†] 김 형 주 ^{**}
 (Sangwon Yoo) (Hongrae Lee) (Hyungdong Lee) (Hyoung-Joo Kim)

요 약 인터넷에 이어 TV환경에서도 넘쳐나는 정보로 인해 시청자들이 원하는 정보를 얻는데 어려움을 겪게 되는 상황이 발생하고 있다. 최근 들어 디지털 방송시대가 열리면서 국내에서도 선택 가능한 채널이 세 자리 수까지 늘어나게 되어 기존의 방법으로는 채널 선택이 곤란하게 되었다. 본 논문은 이러한 문제점의 해결을 위한 TV프로그램 추천 시스템의 프로토타입을 제안한다. 본 논문에서는 TV환경에 알맞은 사용자 정보를 모델링하고 내용기반 방식을 이용하여 추천 시스템의 각 모듈을 설계하였다. 우리의 시스템은 사용자의 기본 정보를 입력받은 후 사용자의 시청 및 사용기록을 추적하여 자동으로 선호도를 파악하고 낱별로 프로그램을 추천한다. 본 논문은 TV 프로그램 추천이 가지는 문제와 이에 대한 접근 방법을 시스템의 구조에 대한 설명과 실험을 통해 밝힌다.

키워드 : 개인화, 추천시스템, TV

Abstract The rapid increase of the number of channels makes it hard to find wanted programs from TV. In recent years, the number of channels come up to hundreds with the digital TV arrival. So, it will drive us to the new way of watching TV. In this paper, we introduce a recommendation system for TV programs to overcome this difficulty. We model user profiles and design each module of the system, considering TV environment. Our system gathers basic information from people manually and then updates user profiles automatically by tracking viewing and usage history. As a result, our system recommends daily TV programs based on the changing interest of users. In this paper, we address the problems and solutions by describing our system and the experiment.

Key words : personalization, recommendation, TV

1. 서 론

TV환경에 새로운 기술들이 접목되면서 TV시청 패턴에도 변화가 불가피하게 되었다. 기존의 공중파 중심으로 이루어진 TV에서는 시청자가 원하는 프로그램을 시청하기 위해 특별히 고려해야 할 사항이 없었다. 케이블 TV를 포함해도 채널과 프로그램의 수가 그리 많지 않기 때문에 본인이 선호하는 채널의 번호를 암기하거나 특정 시간대에 원하는 프로그램을 찾아서 시청하는 일에 별다른 어려움이 없었다.

하지만 디지털 방송시대가 열리고 각 가정마다 디지

털 방송이 일반화되면 채널의 수가 수백 개로 늘어난 상황에서 이전과 같은 채널과 프로그램의 선택은 불가능해진다. 예를 들어 시청 가능한 TV채널이 300개 정도라고 가정해 보자. 이런 상황에서 지금처럼 리모콘을 돌려가며 각 채널에서 원하는 프로그램을 찾으려 한다면 채널당 1초가 걸린다고 해도 300초, 즉 5분의 시간이 걸리게 되는 것이다.

따라서 본 논문은 TV프로그램을 위한 추천시스템을 제안한다. 본 논문에서 구현한 시스템은 시청자로 하여금 모든 채널을 뒤지지 않아도 본인의 취향에 맞는 프로그램을 쉽게 선택할 수 있도록 낱별로 프로그램을 추천해 준다.

TV환경이 아닌 다른 환경에서의 추천내지 필터링 관련 시스템에 관한 연구는 아주 많이 이루어져 왔다[1,2]. 인터넷과 더불어 온라인상의 정보가 기하급수적으로 증가하고 사용자가 원하는 정보를 찾기가 점점 어려워지면서 사용자가 원하는 정보를 직접 찾지 않아도 받아보는 필터링 방식이나 추천 방식이 필요하게 된 것이다.

· 본 논문은 BK21사업 및 ITRC 사업의 지원을 받았음

[†] 학생회원 : 서울대학교 전기.컴퓨터공학부
 swyoo@oopsla.snu.ac.kr
 hrlee@oopsla.snu.ac.kr
 lee@oopsla.snu.ac.kr

^{**} 중신회원 : 서울대학교 전기.컴퓨터공학부 교수
 hjk@oopsla.snu.ac.kr

논문접수 : 2003년 1월 27일

심사완료 : 2003년 6월 9일

추천 시스템을 설계하고 구현함에 있어 기존의 추천 시스템이나 필터링 시스템과 다른 점은 TV라는 환경적인 요인을 충분히 고려해야 한다는 점이다.

첫째, TV는 다른 환경과 달리 TV프로그램을 기술하는 정보가 몇 가지 요소로 대표되기 때문에 사용자 정보 모델링도 이에 맞추어 이루어져야만 한다. 웹 페이지를 대상으로 하는 추천 시스템의 경우 해당 웹 페이지가 가진 정보는 일반적으로 단어들의 집합이며 사용자 정보도 이에 맞추어 단어들의 집합으로 나타난다. 하지만 TV환경의 경우 단어들이 나타난다는 점은 같지만 각 단어가 속한 요소들이 채널, 장르, 출연자, 줄거리, 키워드 등 여러 가지 상위요소로 나누어진다는 점이 다르다. 따라서 이에 맞추어 사용자 정보 또한 모델링 되어야 한다.

둘째, TV환경은 동적이라는 점이다. 웹 페이지의 내용 또한 동적이긴 하지만 그 생명주기가 TV환경처럼 짧지는 않다. TV환경의 경우 영화나 드라마와 같은 경우가 아니면 다시 제공하는 경우도 드물고 프로그램의 가치 자체도 떨어지기 때문에 추천 기법을 선택할 때 주의해야 한다.

현재 추천기법으로 쓰이는 양대 축은 내용 기반 추천(Content-based recommendation)과 협업 기반 추천(Collaborative recommendation)이다. 협업 기반 추천이 실제로 많이 쓰이고 있으며 두 가지 방식을 혼합하여 사용하기도 한다[3,4].

협업 기반 추천은 자신과 취향이 비슷한 사람이 이미 만족스럽다고 평가를 내린 대상을 추천해 주는 것이다. TV환경이 앞으로 양방향 TV시스템으로 이루어져서 사용자가 각 프로그램에 대한 평가를 내릴 수 있겠지만 어떤 대상에 대한 추천이 취향에 비슷한 다른 사람에게 반영되기 위해서는 시간과 관련된 문제가 해결되어야만 할 것이다.

내용 기반 추천은 사용자의 선호도를 표현한 사용자 정보와 추천의 대상이 되는 문서나 아이템의 정보를 비교하여 유사도에 따라 유사도가 높은 문서나 아이템을 추천해 주는 것이다.

셋째, 추천과 관련된 시스템에는 추천과 관련된 수학적 모델이 잘 정립되어 있지 않다. 정보 검색 분야의 경우에는 검색과 관련된 여러 가지 수리적인 모델과 각 모델을 실험할 수 있는 테스트 집합이 잘 이루어져 있지만[5] 추천 시스템의 경우에는 그 평가가 일반 사용자를 대상으로 하며 도메인에 따라 적용이 다르기 때문에 추천과 관련된 모델을 잘 정립해야 한다.

본 논문에서는 위에서 언급한 세 가지 문제들을 해결하기 위한 프로토타입을 구현하고 실험을 통해 시스템의 성능을 보였다.

논문의 구성은 다음과 같다. 2장에서는 추천과 관련된 배경지식을 3장에서는 사용자 정보의 모델링을 서술하였다. 4장에서는 시스템의 구조와 세부 모듈에 대한 설명을 하고 5장에서는 실험 및 실험 결과를 제시하였다. 6장에서 관련연구와 7장의 결론 및 향후 연구로 끝을 맺는다.

2. 배경지식

추천 기법에는 1장에서 언급한 것처럼 크게 내용 기반 방식과 협업 기반 방식이 있다. 이 두 가지 방식은 각자가 가진 장단점이 있으며 적용 도메인에 따라 두 가지 방식을 적절히 혼합하여 사용하기도 한다.

우리는 시스템을 구현함에 있어 내용 기반 방식을 채택하였다. 내용기반 방식은 정보 검색(Information Retrieval)분야에서 기본적인 아이디어를 가져온 것이라고 할 수 있다. 검색 시스템에서 사용자는 시스템에 질의를 던지고 시스템은 사용자에게 질의의 결과로 문서를 순위화하여 보여준다. 이때 순위를 매기는 기준은 사용자의 질의와 질의의 대상이 되는 문서가 얼마나 유사한가이다. 이때 유사도를 측정하는 방법은 여러 가지가 있지만 기본적으로 문서의 내용 즉 문서 안에 있는 단어들과 질의가 가진 단어들을 비교하는 것으로 이루어진다.

내용 기반 추천기법도 마찬가지로 방식을 취한다. 차이점은 사용자가 원하는 정보를 찾기 위해 질의를 던지기 전에 시스템이 추천을 해야 하므로 사용자의 질의는 사용자 정보(user profile)로 대체된다. 사용자 정보는 개념적으로 장기간 동안 변하지 않는 질의라고 생각할 수 있다. 내용 기반 추천 시스템의 일반적인 구조는 그림 1과 같다.

그림 1을 보면 내용 기반 추천 시스템에서 가장 중요한 두 가지 요소가 사용자 정보와 추천의 대상이 되는 아이템에 대한 정보이다. 사용자 정보는 사용자가 자신의 선호도를 표시한 것들로 이루어진다.

예를 들어 영화에 관한 추천 시스템이라면 사용자의 선호도는 사용자가 좋아하는 배우나 감독, 영화의 장르

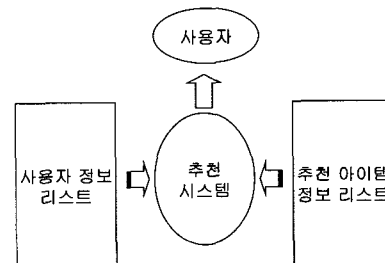


그림 1 내용 기반 추천 시스템의 일반적인 구조

등으로 이루어 질 수 있다. 추천의 대상이 되는 영화에 관한 정보 또한 배우, 감독, 장르 등으로 이루어져 있어야 해당 요소간의 비교 및 유사도 측정이 가능해진다.

내용기반 추천이 가진 단점은 다음과 같다. 만약 추천 대상이 되는 아이템에 관한 정보가 부족하다면 추천의 질이 떨어질 수 밖에 없다는 것이다. 예를 들어 영화라면 해당 영화가 어느 나라에서 만든 영화인지 알아야 영화의 국적에 대한 선호도에 대응하여 추천이 가능해진다. 둘째, 사용자는 자신이 가진 취향의 테두리 안에 갇힐 우려가 있다는 것이다. 내용 기반 추천 시스템의 경우 사용자 정보에 나타난 사용자의 선호도에 근거하여 추천을 한다. 어떤 사용자가 액션영화를 좋아한다고 선호도를 밝혔다면 추천이 되는 영화 중 상위에 랭크되어 있는 영화는 대부분 액션영화일 가능성이 많다. 가끔씩 이 사용자가 보고 싶은 영화가 액션이 아닌 멜로물이라고 해도 내용 기반 방식으로는 추천될 가능성이 낮아진다. 마지막으로 사용자가 의사 표현을 얼마나 많이 하느냐에 따라 추천의 질이 달라진다는 것이다. 사용자가 추천한 내용에 대해 평가를 내리면 시스템은 이를 바탕으로 사용자 정보를 수정할 수 있다. 사용자의 의사 표현은 많으면 많을수록 시스템이 정확한 추천을 하는데 도움이 되지만 사용자에게는 부담스러운 일일 수 있다.

협업 기반 추천 시스템은 취향이 비슷한 사람을 하나로 묶어 그 그룹 내에서 다른 사람들이 본 영화라면 나에게도 흥미가 있을 것이라는 가정에서 출발한 방식이다. 사람의 평가 또는 추천에 근거를 두고 있으므로 추천의 대상이 되는 아이템에 관한 정보가 부족하거나 기술하기 곤란한 경우에도 적용 가능하다. 마찬가지로 이유로 자신이 선호한다고 입력한 것과 내용 기반 방식에 의하면 거리가 있는 것들도 그룹 내의 다른 사람이 추천할 경우 나에게 추천될 수 있다.

협업 추천 방식이 가진 단점은 상대적으로 내용 기반 방식의 장점으로 이해될 수 있다. 첫째, 자신이 가진 취향과 비슷한 사람이 적을 경우 추천의 질이 떨어질 수 있다. 자신과 비슷한 선호도를 가진 누군가가 평가를 해준 아이템을 추천 받아야 하기 때문에 비슷한 사람의 수가 적으면 상대적으로 손해를 볼 수밖에 없다. 둘째, 추천의 대상이 새로 주어지는 경우 이에 대해 누군가가 평가를 하기 전에는 추천이 이루어지지 않게 된다. 예를 들어 새로운 영화가 개봉했을 때 협업 기반 추천 시스템은 그 영화에 대한 누군가의 평가가 있기 전까지는 그 영화를 추천해 주지 않는다. 만약 자신과 비슷한 취향을 가진 그룹 내의 누구도 그 영화를 보지 않고 망설이고 있는 중이라면 그 영화는 계속 추천되지 못하고 남게 된다.

TV프로그램을 위한 추천시스템의 경우 두 가지 기법을 택함에 있어 두 가지 기법이 가진 기본적인 장단점 이외에도 환경적인 요인을 고려해야만 한다. 협업 기반 방식이 우수함에도 불구하고 적용하기 곤란한 이유는 다음과 같다. 첫째, 협업 기반 방식의 경우 다른 사람이 추천한 프로그램을 내가 시청해야 하는데 어떤 프로그램을 시청한 후에 추천을 해 준다면 이미 그 프로그램은 더 이상 볼 수 없게 된다. 시간적인 문제가 생기게 되는 것이다. VOD 환경의 경우에도 영화나 드라마와 같은 특정 장르가 아니면 일반적으로 서비스가 되지 않고 일정시간이 지나면 뉴스와 같은 장르는 의미가 없어지기 때문에 적절하지 않다. 둘째, 협업 기반의 경우 많은 시청자들의 시청기록과 프로그램에 대한 평가들을 관리해 줄 중앙 서버가 필요하다. 반면 내용 기반 방식의 경우 현재 상용화 되어 있는 것처럼 TV프로그램 정보만 정기적으로 제공 받으면 쉽게 셋탑박스와 간단한 저장장치 만으로도 구현 가능하다.

본 논문에서 구현한 추천 시스템은 내용 기반 방식을 택하였다. 물론 추후에 여러 가지 환경적인 요인들이 바뀐다면 협업 추천 방식도 도입 가능할 것이다.

3. 사용자 정보 모델링

내용 기반 추천에서 사용자 정보 모델링은 추천의 대상이 되는 아이템에서 얻어낼 수 있는 정보와 밀접하게 연관을 맺고 있다. 본 논문은 추천의 대상이 되는 프로그램 정보가 TVAnytime[6]의 표준을 따른다고 가정하였다. TVAnytime 포럼은 1999년부터 활동을 시작하였으며 미국, 유럽, 아시아의 방송국 및 방송기기 제조업체들이 참여하여 차세대 TV환경과 관련된 여러 가지 표준을 정하는 조직이다. 이 포럼에서 발표한 표준에는 방송 프로그램의 메타데이터와 관련된 규약도 있으며 XML 스키마로 이를 기술하고 있다. 따라서 우리는 이 표준에 대응하는 시청자의 선호도를 다음의 요소들로 모델링 하였다.

- 장르, 세부장르 : 장르와 세부장르는 장르구분에 계층적인 구조가 존재함을 나타낸다. 장르에 뉴스가 존재한다면 세부장르는 스포츠 뉴스, 날씨 뉴스, 경제 뉴스 등이 존재할 수 있다. 장르와 세부장르라는 계층적인 구조를 둔 이유는 방송 프로그램의 장르 구분이 영화와는 다른 특징이 있기 때문이다. 영화 내에서만 장르 분류를 한다면 일단 형식은 영화로 정해지고 내용에 따라 장르를 나누는 것이 일반적이다. 하지만 방송 프로그램의 경우는 먼저 형식에 따른 분류가 필요하므로 계층적인 구조를 두었다. 따라서 시청자의 선호도가 형식적인 것을 따른다면 장르 분류만으로도 충분할 것이고 시청자가 내용 중심의 선호도를 가지고 있으면 세부 장르에 대한

선호도까지 필요하게 된다.

예를 들어 어떤 사람이 다큐멘터리를 좋아한다면 이 사람의 선호도는 “장르-다큐멘터리”로 나타나고 그중에서도 환경문제에 관심이 있다면 선호도는 “장르-다큐멘터리 세부장르-환경”으로 나타나게 된다.

장르와 세부장르는 일정한 범위의 집합으로 표현 가능하고 장르의 구분 자체는 주제에 따라 주관적일 수 있다. 이 역시 TVAnytime에서 주어진 분류체계를 참고하였다. 본 논문에서는 시청자가 세부장르에 대한 선호도를 가질 경우 장르에 대한 선호도도 가지며 그 역은 성립하지 않는다고 가정하였다. 장르와 세부 장르에 대한 선호도는 복수일 수 있다.

• 채널 : 채널은 TV프로그램이 방영되는 채널을 의미한다. 앞으로 채널의 수가 많아지면 많아질수록 채널 별 특성이 두드러지게 나타날 것이다. 예를 들어 24시간 영화를 내보내는 채널 A와 B가 있다고 하자. 장르의 분류상으로는 이들은 똑같은 영화로 구분이 가지 않는다. 하지만 사람들마다 재미있다, 재미없다 또는 유려다, 무르다 등으로 채널에 대한 선호도가 정해질 수 있다.

채널의 숫자가 한정되어 있으므로 일정한 범위의 집합으로 표현 가능하고 시청자는 복수의 채널에 대해 선호도를 가질 수 있다.

• 출연자 : TV프로그램의 출연하는 사람들을 말한다. 좀 더 포괄적인 개념으로 보면 TV프로그램을 만든 스텝들로 PD나 엔지니어, 작가 등까지도 포함될 수 있지만 방송에 직접 출연하는 사람들로 한정하였다. 출연자의 범위는 매우 다양해서 일정한 범위의 집합으로 표현 불가능하다. 출연자는 배우, 탤런트, 가수, 운동선수, 팀, 아나운서 등이 될 수 있다. 채널이 많아지고 시청자들의 취향이 다양해지면 카레이싱 선수나 프로 게이머와 같은 출연자들도 프로그램에 대한 선호도를 결정짓는 중요한 요인이 될 것이다. 예를 들어 같은 뉴스라도 시청자가 선호도를 표시한 앵커가 있다면 그 뉴스가 추천될 수 있다.

• 줄거리: 줄거리는 실제 방송되는 프로그램의 내용을 담고 있다. 시청자의 줄거리에 대한 선호도는 방송내용에 관한 단어들로 표현될 수 있다. 줄거리에는 장르별로 다양한 내용이 포함될 수 있다. 요리 채널의 프로그램이 “장르-취미 세부장르-요리”라고 분류된다면 그날의 요리가 무엇이나에 따라 방송 내용은 달라지게 된다. 시청자는 중식, 한식, 일식 중 어느 하나에만 선호도를 가질 수 있기 때문에 줄거리를 통해 원하는 프로그램을 추천해 줄 수 있다. 줄거리는 일정한 범위로 표현 불가능하다.

• 키워드(주제어) : 마지막으로 각 방송 프로그램은 주제어에 의해 표현될 수 있다. 이는 특정 장르의 경우 장르나 줄거리만으로 표현할 수 없는 주제가 있을 수

있기 때문이다. 예를 들어 전쟁이나 사랑, 우정, 가족 등이 그 예가 될 수 있고 시청자는 이러한 추상적인 선호도를 가질 수 있다. 주제어는 일정한 범위로 표현 불가능하다.

본 논문에서 위와 같이 TV환경을 고려하여 시청자의 선호도 정보를 다섯 가지로 모델링 하고 각각이 가중치를 가질 수 있도록 하였다. 이때 가중치는 두 가지 형태로 나타나는데 첫째는 각 요소의 값들에 대한 가중치이다. 예를 들어 내가 어떤 채널을 좋아한다고 하면 “A라는 채널을 1만큼 좋아할 때 B라는 채널을 3만큼 좋아한다.”라고 표현할 수 있다. 두 번째는 각 요소간의 반영 비율이다. 어떤 시청자는 시청 패턴이 출연자에 의해 크게 좌우될 수 있고 또 다른 시청자는 시청 패턴이 채널에 의해 크게 좌우될 수도 있다. 이를 표현하기 위해 전체 선호도를 100이라고 할 때 각 시청자는 각 요소별 반영 비율에 대한 선호도를 가질 수 있다 이를 그림으로 표현하면 그림 2와 같다. 수학적인 모델과 프로그램의 순위화 방법은 4절에서 설명한다.

사용자 정보	가중치	반영비율
장르	드라마 1, 스포츠 4	10%
세부장르	드라마,액션 3 스포츠,야구2	20%
채널	KBS 3, MBC 4	5%
출연자	김두한 5, LG트윈스 2	30%
줄거리	겨울 1, 월세 2	5%
주제어	사랑 3,삼각관계 3,우정 1	30%

그림 2 사용자 정보 모델링

4. 시스템 구조

이 절에서는 우리가 구현한 시스템과 각 세부 모듈 그리고 매칭 기법에 대해 설명한다. 우리가 구현한 시스템의 전체적인 구조는 그림 3과 같다. 양방향 TV를 시뮬레이션하기 위해 웹 인터페이스를 사용하였으며 JAVA와 JSP를 이용하여 구현하였다. 사용자 정보의 저장은 데이터베이스의 테이블을 통해 이루어진다.

전체적인 워크 플로우를 살펴보면 사용자 정보는 사

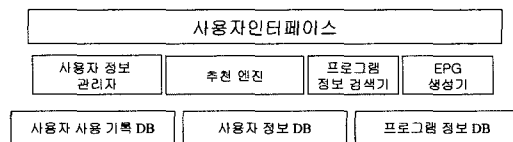


그림 3 시스템의 전체적인 구조

용자 정보 관리자에 의해 관리된다. 사용자 정보 관리자는 시스템 사용자로부터 해당 사용자를 식별할 수 있는 아이디와 사용자의 초기 선호도를 얻는다. 사용자의 선호도는 사용자 정보 DB에서 관리되며 각 요소별로 별도의 테이블을 가진다. 사용자 정보 관리자는 사용자의 추천에 대한 평가를 사용자 정보에 반영하여 사용자 정보를 수정하는 역할을 한다. 또 시스템 사용자는 이 시스템에서 자신이 원하는 프로그램을 검색하거나 프로그램 정보를 열람할 수 있으며 이에 대한 기록은 사용자 사용기록 DB에 저장된다. 사용자 정보 관리자는 이를 분석하여 사용자 정보에 반영한다.

추천엔진은 사용자 정보DB와 프로그램 정보DB의 내용을 비교하여 해당 사용자의 선호도와 각 프로그램 사이의 유사도를 계산해낸다. 이 유사도에 의해 프로그램은 순위화되고 사용자에게 추천이 이루어진다.

프로그램 정보DB는 3절에서 밝힌 것처럼 TVAny-time표준을 따르는 XML데이터를 파싱하여 DB테이블의 형태로 관리한다.

프로그램 정보 검색기는 사용자가 자신이 원하는 프로그램을 검색할 수 있도록 지원해 준다. 사용자는 자신이 원하는 프로그램을 장르, 인물, 채널, 주제어 등을 이용해 일자별로 검색할 수 있다.

EPG는 Electronic Program Guide의 약자로서 오프라인 상의 편성표를 화면으로 옮겨놓은 것을 말한다. 시스템은 프로그램 정보를 이용하여 사용자에게 자신이 원하는 프로그램에 관한 정보를 얻을 수 있는 인터페이스를 제공한다.

4.1 사용자 정보 관리자

사용자 정보 관리자의 핵심 기능은 사용자로부터 선호도에 관한 초기값을 얻어내는 일과 이후 사용자의 시청기록, 추천 프로그램에 대한 평가를 지속적으로 사용자 정보에 반영해 나가는 일이다.

사용자 정보의 초기값을 얻어내는 일은 한번만 수행하면 되는 작업이지만 사용자에게 여러 가지를 요구할 경우 사용자로부터 적절한 정보를 얻어 내는데 실패할 수 있다. 본 논문에서 구현한 시스템은 사용자 정보를 모델링하기 위해 장르, 세부장르의 계층적인 구조를 사

용하고 있다. 장르당 세부장르의 수가 상당히 많기 때문에 이를 좀 더 효과적으로 입력 받기 위해 대표 프로그램을 통한 입력 방식을 취하였다. 대표 프로그램을 통한 입력 방식이란 일단 사용자로부터 선택의 수가 적은 장르에 대한 선호도를 입력 받은 뒤 이를 이용해 시스템이 세부 장르에 해당하는 대표 프로그램을 생성해 사용자의 쉬운 선택을 유도하는 것이다.

예를 들면 사용자가 드라마를 선호한다고 입력하면 시스템은 그림 4의 오른쪽에 나타난 것처럼 사용자에게 드라마의 세부장르를 대표하는 프로그램들을 보여준다. 이때 “야인시대”는 액션, “겨울연가”는 로맨스를 대표하게 된다. 이런 방식을 취하게 되면 분류 체계의 이름만으로는 확실하지 않은 프로그램의 성격이 쉽게 드러나기 때문에 사용자는 좀 더 자신의 선호도를 쉽게 표현할 수 있다. 대표 프로그램의 선정은 주관적일 수 있기 때문에 시청률 데이터를 이용한다고 가정하였다. 시청률 데이터중 상위 랭크된 프로그램들을 제시하여 사용자의 인지도를 제고하고 대표성을 가지는 TV프로그램의 시간에 따른 변화를 반영하였다.

그림 4의 사용자 평가 처리기는 사용자의 프로그램 추천에 대한 평가를 반영한다. 사용자의 추천 프로그램에 대한 평가는 가장 확실한 의사 표현이므로 즉시 사용자 정보에 반영한다. 추천에 대한 평가는 사용자의 편의를 위해 긍정과 부정, 2단계 스케일을 사용하였다.

$$W_{i+1} = \alpha W_i + \beta \sum_{d \in R} v_d - \gamma \sum_{d \in NR} v_d \quad (1)$$

식 (1)은 사용자 정보관리자가 추천에 대한 평가를 그림 2에 나타난 가중치에 반영하는 식이다. W_i 는 현재 사용자 정보가 가진 가중치를 나타내며 W_{i+1} 은 수정후의 사용자 정보가 가진 가중치를 나타낸다. W_i 는 벡터와 같이 표현된다. 예를 들어 장르에 대한 가중치라면 $W_i=(3, 0, 1, \dots)$ 과 같이 벡터의 각 차원이 드라마나 오락, 뉴스등과 같은 요소들에 대한 가중치 정보를 표현한다. v_d 는 해당요소의 차원에 값이 존재하면 1 아니면 0을 가진다. R 은 사용자가 긍정적 반응을 보인 경우이고 NR 은 사용자가 부정적 반응을 보인 경우이다. 가중

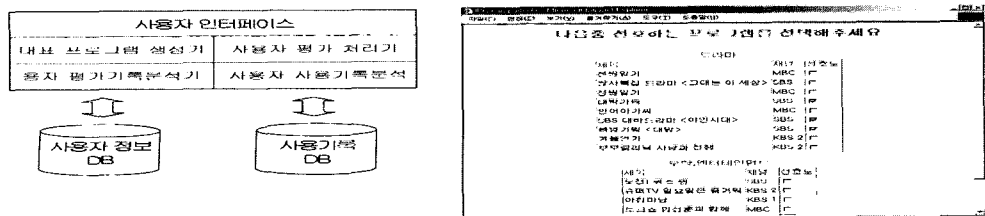


그림 4 사용자 정보 관리자의 구조와 대표프로그램 생성기

치 W 는 그림 2에 나타난 것처럼 장르, 세부장르, 채널 등에 대해 각각의 벡터 값을 가지게 된다. α, β, γ 는 계수로써 가중치의 변화율을 조정해주는 역할을 한다. 기존 사용자 정보에 없는 항목의 경우는 새로 추가해 주고 가중치는 1로 초기화한다.

그림 4의 사용자 평가 기록 분석기는 추천을 할 때 장르, 서브장르, 채널 등의 반영 비율을 계산해 내기 위한 부분이다. 이를 도입하는데 바탕이 되는 아이디어는 3장에서 언급한 것처럼 사용자 별로 선호도의 주된 대상이 다를 수 있다는 것이다. 이를 알아내기 위해 사용자의 추천 프로그램에 대한 평가를 이용하였다.

사용자 평가 처리기는 사용자가 원하는 추천을 얻기 위한 1단계 작업을 수행하고 사용자 평가기록 분석기는 2단계 조정이라 할 수 있다. 추천된 프로그램에 대해 사용자가 내린 평가가 누적되면 사용자가 원하는 프로그램에 대한 추천비율이 처음보다 높아지게 된다. 즉 액션 영화를 좋아하는 사용자에게 다큐멘터리가 추천되지는 않는 것이다. 하지만 사용자에게 영화채널에 대한 선호도가 있다면 이로 인해 특정 채널의 영화를 추천해야 할 필요가 있을 수 있다. 채널, 장르, 출연자 등에 모든 요소에 고른 취향을 가진 사용자라면 여기서 추가 조정의 필요성이 없는 것이고 이 중 한쪽에 쓸림이 있는 사용자라면 추가 조정이 필요하게 된다.

평가기록 분석기는 식 (2)에 따라 사용자가 평가를 내린 프로그램들을 긍정적 평가와 부정적 평가의 두 부류로 나누고 이 기록을 누적하였다가 일정 주기로 분석한다.

$$\text{Sum}_i = \sum_{j=1}^n \alpha * \text{positive_cnt}(j) + \beta * \text{negative_cnt}(j)$$

$$\text{Modify_ratio}(\text{Sum}_1, \text{Sum}_2, \dots, \text{Sum}_i, \gamma) \quad (2)$$

식 (2)의 Sum은 n개의 프로그램에 대한 장르, 서브장르, 채널, 출연자, 키워드 각각에 대한 긍정적 평가와 부정적 평가 횟수의 합을 나타낸다. positive_cnt는 긍정적 평가의 횟수이며 negative_cnt는 부정적 평가의 횟수이다. α 와 β 는 계수를 나타낸다. 예를 들어 시스템이 어떤 프로그램을 추천한 근거가 채널인데 사용자가 긍정을 표시하면 채널에 대한 Sum은 증가하고 부정을 표시하면 Sum은 감소한다. Modify_ratio는 Sum들과 조정비율을 인자로 하여 일정 주기로 반영비율을 조정해 준다.

식 2에서 α 는 양의 값을 β 는 음의 값을 부여한다. 여기서 차이를 구하는 이유는 사용자가 긍정을 표시했다는 것은 추천의 원인이 된 요소들이 사용자에게 영향을 끼쳤다는 것이고 부정을 표시한 것은 영향을 끼치지 못했다는 것으로 해석할 수 있기 때문이다.

각 요소별 무게가 구해지면 반영비율의 조정이 이루어진다. 수정 비율의 크기가 10%인 경우 채널의 무게가 +10이고 출연자의 무게가 -10이며 다른 요소는 0이라면 채널은 5%반영 비율이 증가하게 되고 출연자는 5% 감소하게 된다.

결국 해당 프로그램의 추천 근거가 출연자인 경우 사용자가 이를 받아들이면 인물에 대한 의존도를 높이고 사용자가 관심을 보이지 않으면 출연자의 비중을 낮추게 된다.

그림 4의 사용자 사용기록 분석은 두 가지로 나누어지며 식 (3)에 의해 사용자 정보에 반영이 이루어진다. 먼저 검색기록에 대한 분석이다. 채널이 수백 개에 이르게 되면 하루에 방영되는 프로그램의 개수가 만여 개에 육박하게 된다. 이런 상황에서 자신이 원하는 프로그램을 직접 찾으려면 TV에서 제공하는 검색 기능을 이용할 수밖에 없다.

사용자가 이용할 수 있는 검색의 대상은 프로그램 제목, 채널, 일자, 출연자, 장르, 세부장르, 키워드이다. 특정 항목에 대한 검색은 사용자의 요구를 반영하는 것으로 사용자의 추천에 대한 평가와 마찬가지로 직접적인 의사표시로 파악할 수 있다.

$$W_{i+1} = \alpha W_i + \beta \sum_{d \in S} v_d \quad (3)$$

식 (3)은 식 (1)과 마찬가지로 W_i 는 현재 사용자 정보가 가진 가중치 W_{i+1} 은 수정후의 사용자 정보가 가진 가중치를 나타낸다. v_d 는 장르, 세부장르, 채널, 인물 등의 차원을 나타낸다. S는 사용자가 검색을 한 값들이며 α, β 는 계수이다. 식 (1)과의 차이점은 사용자가 검색한 각 항목은 사용자 정보에 반영되어 가중치를 증가시키지만 하며 감소시키는 경우는 없다는 것이다.

두 번째로 사용자의 시청기록이다. 본 시스템에서는 EPG를 이용하여 해당 일자에 TV를 시청한 기록을 얻어내는 것으로 실험하였다. 직접 TV를 본 시청데이터를 얻는 것이 불가능하여 EPG에 프로그램과 관련된 정보를 링크로 처리하고 사용자로 하여금 해당 프로그램의 시청을 원할 경우 시청한다는 의사표현을 할 수 있도록 하였다. 실제 TV환경에서는 시청기록 중 일정시간 이상의 것만을 선택해야 하지만 실험에서는 사용자가 끝까지 프로그램을 시청했다고 가정한다.

따라서 시청기록으로부터 얻어낸 프로그램 리스트로부터 각 요소를 추출하여 식 (3)을 이용 사용자 정보를 수정해 준다.

4.2 추천 엔진

추천 엔진은 사용자 정보와 프로그램 정보를 검색하여 유사성을 비교한 후 프로그램에 순위를 부여하는 역

할을 한다. 3절에서 언급한 것처럼 사용자는 2가지 종류의 선호도를 가지게 되는데 하나는 장르나 채널과 같은 각 요소의 값들에 대한 선호도이며 다른 하나는 자신이 여러 요소 중 어느 쪽에 더 영향을 많이 받는가에 대한 선호도이다.

사용자의 요소 값들에 대한 선호도는 양의 값과 음의 값의 일정한 범위 내에서 가지며 초기 정보 입력시 사용자로부터 얻은 값들은 1로 초기화된다. 사용자의 요소별 영향 비율에 대한 반영은 균등하게 반영되는 것으로 초기화 된다.

선호도에 대한 유사도 측정은 두 가지로 나뉜다. 채널, 장르, 세부장르와 같이 프로그램 정보가 단일한 값을 가지는 경우와 출연자, 줄거리, 키워드와 같이 복수의 값을 가질 수 있는 경우이다. 단일한 값을 가지는 경우는 일치한 값이 가진 가중치가 해당 요소의 유사도가 된다. 복수의 값을 가지는 경우 출연자와 다시 줄거리, 키워드로 구분하였다. 출연자의 경우 사용자 정보와 하나의 프로그램 정보에서 복수의 출연자가 일치할 경우 더 많은 가중치를 줄 것인가 하는 문제는 곤란한 문제이다. 자신이 좋아하는 배우나 탤런트가 2명 이상이 함께 출연한다고 해서 해당 프로그램을 더 좋아하게 될 것이라고 보장할 수 없기 때문이다. 따라서 본 시스템에서는 사용자의 출연자 선호도중 가장 가중치가 높은 값을 택하였다. 줄거리와 키워드는 정보검색 분야에서 사용되는 벡터 모델의 tf-idf 기법[5]을 이용하여 유사도가 계산 가능하다. 각 프로그램이 가진 사용자 정보와의 유사도는 요소별 가중치가 나오고 여기에 다시 요소별 반영비율을 곱하여 최종 결과가 산출된다.

사용자에게 시스템이 추천한 결과가 그림 5에 나타나 있다. 순위에 따라 추천 프로그램을 보여주고 점수란에 약어 표시가 나타나는데 G는 장르 S는 세부장르 C는 채널 등으로 추천된 프로그램이 어떤 근거에서 추천이

되었는지를 사용자에게 알려주어 추천된 프로그램을 선택하여 시청하는데 도움을 주고 있다.

시스템이 추천한 결과 중 같은 시간대의 프로그램이 있는 경우도 있는데 시청자입장에서 모두를 시청하고 싶다면 시청을 원하는 프로그램을 녹화하면 되므로 추천시에 이들 모두를 보여주는 방식을 택하였다.

이외에 본 시스템은 XML로 이루어진 프로그램 정보 메타데이터를 DB로 저장하는 기능과 이 정보를 바탕으로 EPG화면을 생성해 내는 기능, 프로그램 검색 기능을 가지고 사용자의 TV프로그램 시청을 도와준다.

5. 실험 및 결과

TV프로그램 추천을 위한 실험은 환경의 제약으로 인해 8일분 지상파 5개 채널(KBS1, KBS2, MBC, SBS, EBS) 데이터를 이용하여 이루어졌다. 본 논문에서 가정한 환경을 직접 실험하기 위해서는 수백 개 채널의 수개월 분에 해당하는 데이터를 필요로 하지만 현재 이를 제공할 수 있는 기관이 없어 8일 분의 데이터만으로 가능성을 보이도록 하였다.

TV프로그램을 기술하는 프로그램 정보는 각 방송사에서 제공하는 편성표를 기준으로 만들어졌고 각 방송사에서 제공하는 데이터의 부족으로 인해 실험상에서 시스템의 기능을 수정하여 사용한 것은 다음과 같다.

먼저 사용자 정보 중 채널에 대한 선호도를 배제하였다. 지상파는 한 채널이 한 주제가 아닌 다양한 내용으로 구성되므로 채널에 대한 선호도가 나타나기에 곤란하다는 판단을 하여 이를 제외하였다. 둘째로 줄거리에 관한 선호도를 배제하였다. 프로그램 데이터의 내용을 분석해 본 결과 드라마나 영화를 제외한 다른 형식의 경우 프로그램의 내용에 대한 기술이 부실한 경우가 많아 이를 사용하기 어려웠다. 셋째, 채널의 개수가 5개에 불과하여 프로그램 정보가 한눈에 들어올 수 있기에 실험결과 검색기능의 사용이 미미하여 이를 제외하고 사용자의 시청기록만 식 (3)을 이용하여 반영하였다.

실험에 사용한 프로그램의 총수는 1036개이며 8명의 사람이 실험에 참여하였다. 실험 참여 방법은 다음과 같다. 먼저 시스템에 접속하여 자신의 간단한 선호 정보를 초기화 한다. 이후 7일 동안 시스템이 추천해 준 프로그램에 대해 긍정과 부정 두 가지 평가를 내리는 것과 실제 시청한 프로그램을 입력하는 것으로 이루어졌다. 사용자의 평가 스케일은 일반적으로 3가지 스케일이나(긍정, 부정, 보통) 5가지 스케일(아주 좋음, 좋음, 보통, 나쁨, 매우 나쁨)도 사용하며 추천 결과를 더 좋게 만들 수 있다. 그러나 실제 TV에서 사용자가 자세한 평가를 내리기는 어렵다고 보고 긍정과 부정 두 가지 스케일로 사용자의 평가를 입력 받았다.

그림 5 추천 결과 페이지

시스템의 성능 척도는 coverage를 사용한다. coverage는 일자별로 사용자가 시청한 프로그램들이 얼마나 많이 추천 리스트에 포함되어 있는나에 관한 척도이다 [7]. 추천 시스템을 평가하는 척도로 영화와 관련된 분야에서는 MAE(Mean Absolute Error)가 많이 사용된다[8]. 하지만 TV와 같이 사용자 입장에서 프로그램을 본다 안본다의 두 가지로만 나누는 경우는 IR의 precision[7]과 유사한 coverage가 더 적절하다고 볼 수 있다.

$$coverage\ ratio = \frac{n(R \cap W)}{n(W)} \quad (4)$$

식 (4)는 coverage의 정의를 나타낸 식이다. R은 시스템이 추천한 리스트를 나타내며 W는 실제 사용자가 시청한 리스트를 나타낸다.

표 1에 실험결과가 정리되어 있다. 8명의 사용자는 7일 분의 프로그램에 대해 시스템의 추천에 대한 평가를 내렸으며 7일 분의 방송에 대해 시청한 프로그램을 기록하였다.

시스템의 성능에 대한 평가는 마지막 8일째 방송에 대한 시스템의 추천과 사용자의 시청기록을 비교하는 것으로 이루어졌다. 8일째 방송분은 134개의 프로그램으로 이루어져 있고 이 중 사용자가 시청한 프로그램의 숫자는 초기 프로파일 항목 옆 괄호 안에 쓰여진 숫자이다.

1번 사용자를 예로 들어 설명하면 이 사용자는 7일 동안 39회의 추천에 대한 평가를 내렸으며 초기 정보만으로 시스템이 추천한 결과의 실제 사용자가 시청한 프로그램에 대한 coverage는 0.25와 0.41이다. 0.25는 시스템이 추천한 프로그램의 상위 10%와 비교한 경우이며 0.41은 상위 25%와 비교한 경우이다. 그 다음 시스템이 사용자의 추천에 대한 평가를 반영하도록 했다. 두 줄의 숫자쌍 중 위 줄은 4일 분을 반영한 것이며 아래 줄은 3일을 더한 7일분을 반영한 것이다. 4일분을 반영한 경우 상위 10%에서 coverage는 0.08이며 상위 25%

에서는 0.5이고 7일분을 반영하면 상위 10%에서 0.08, 상위 25%에서 0.41이다.

그 다음 항목은 시청 횟수인데 1번 사용자는 7일간 총 80개의 프로그램을 시청했음을 나타낸다. 이러한 시청기록을 4일분과 7일분씩 반영하면 각각 상위 10%에서 coverage는 0.41에서 0.50으로 상위 25%에서 coverage는 0.58에서 0.58로 나타난다.

실험 결과에 나타난 계수들은 모두 1로 고정하였다. 정보 검색 분야에서 사용자의 피드백을 반영할 때 적절한 계수에 관한 대표적인 실험은 [9]에 나타나 있으며 적절한 계수는 $\alpha=1, \beta=2, \gamma=0.5$ 로 알려져 있다.

하지만 이 값은 TV가 아닌 문서의 경우 이므로 본 실험에서 계수의 값을 조정해가며 실험해 본 결과 $\alpha=1, \beta=1, \gamma=1$ 일 때 사용자의 피드백을 적절히 반영하는 것으로 나타났다. 이 계수의 의미는 긍정과 부정을 같은 비율로 반영해 주는 것이다. 하지만 실험 참여자의 수가 적어 통계적으로 의미가 있다고 판단하기에는 곤란하기 때문에 좀 더 정확한 값을 찾아내려면 다수의 실험 참여자가 필요하다.

식 (2)를 이용하여 평가기록을 누적한 후 사용자별로 추천의 대상이 되는 요소간 반영 비율을 조정해 주는 실험은 일부 사용자에게 값의 변화가 나타났으나 표 1에는 나타나지 않았다. 그 원인은 실험기간과 참여인원의 부족에도 원인이 있겠지만 채널별로 콘텐츠의 전문화가 이루어지 않았기 때문이라고 판단된다. 실험에 사용한 지상파 방송의 성격상 종합 편성이 이루어지기 때문에 이런 영향이 잘 나타나지 않는다고 볼 수 있다.

그림 6을 보면 사용자의 초기 정보에서 사용자가 시스템에 추천에 대한 평가와 시청기록을 통해 추가 정보를 제공함으로써 추천의 질이 향상됨을 확인할 수 있다. feedback 2는 사용자의 추천에 대한 평가가 4일분이 반영된 경우이며 feedback 3는 추천에 대한 평가 3일분이 추가로 반영된 경우이다. feedback 4는 추천에 대한 평가를 반영한 후에 시청기록 4일분이 추가로 반영된

표 1 실험 결과

사용자	추천에 대한 평가	초기 프로파일	사용자의 피드백 반영	시청횟수	사용자의 시청기록반영
1	39회	(12)0.25, 0.41	(0.08, 0.50)/(0.08, 0.41)	80회	(0.41, 0.58)/(0.50, 0.58)
2	56회	(2)0.00, 0.00	(0.50, 0.50)/(0.50, 0.50)	31회	(1.00, 1.00)/(1.00, 1.00)
3	57회	(3)0.66,0.66	(0.33, 0.66)/(0.33, 0.66)	45회	(0.66, 1.00)/(1.00, 1.00)
4	63회	(7)0.28,0.28	(0.14, 0.28)/(0.14, 0.28)	70회	(0.42, 0.71)/(0.57, 0.85)
5	74회	(5)0.40, 0.40	(0.00, 0.40)/(0.20, 0.60)	39회	(0.40, 0.40)/(0.60, 0.60)
6	73회	(6)0.33, 0.33	(0.33, 0.33)/(0.50, 0.50)	63회	(0.83, 1.00)/(0.83, 1.00)
7	156회	(9)0.55, 0.77	(0.44, 0.55)/(0.44, 0.66)	46회	(0.55, 0.77)/(0.66, 0.88)
8	176회	(12)0.33, 0.50	(0.41, 0.41)/(0.33, 0.58)	92회	(0.41, 0.66)/(0.41, 0.58)
평균	86회	0.35, 0.41	(0.27, 0.45)/(0.31, 0.52)	49회	(0.58, 0.76)/(0.69, 0.81)

실험에 사용된 계수 (식 1) $\alpha=1, \beta=1, \gamma=1$ (식 3) $\alpha=1, \beta=1$

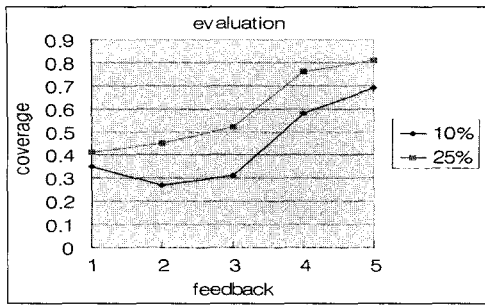


그림 6 사용자 평가와 시청기록에 따른 coverage의 변화

것이며 feedback 5는 나머지 3일 분이 반영되었음을 나타낸다.

coverage가 떨어졌다가 다시 올라가는 현상은 사용자가 자신이 입력한 초기 정보가 어떤 형태로 반영이 될지 정확히 알 수가 없기 때문에 이에 적응해 나가는 기간이라고 분석할 수 있다.

6. 관련 연구

추천과 관련된 연구는 TV가 아닌 다른 환경을 중심으로 많은 연구가 있어 왔다. 추천시스템의 주된 대상은 영화데이터였으며 영화를 대상으로 여러 가지 기법을 테스트하고 그 결과를 보였다. [10]은 협업 기반 방식의 알고리즘들을 실험적으로 분석하였으며 [11]에서는 협업 기반 방식과 내용 기반 방식을 혼합하여 추천의 결과를 향상시키는 방법에 대해 제안하였다.

TV와 관련된 연구는 다음과 같다. [12]는 온라인상에서 프로그램 정보를 제공하고 이를 이용하여 취향이 비슷한 사용자들끼리 묶은 후 CBR(Case Base Reasoning)방식과 협업기반 방식을 혼합한 방식으로 추천리스트를 제공하였다. [13]은 Agent를 기반으로 사용자에게 프로그램을 추천하는 방식을 취하고 있다. 최근 들어 TV의 개인화와 관련한 워크샵이[14] 개최되면서 연구가 본격적으로 시작되고 있다.

상용화되어 있는 서비스로는 다음과 같은 것들이 있다. Canal plus[15]는 1993년부터 양방향 TV서비스를 위한 미들웨어를 공급해온 회사로 유럽 전역을 대상으로 방송사업을 펼치고 있으며 추천은 아니지만 본인이 원하는 정보를 손쉽게 검색하는 기능을 제공한다. Tivo[16]는 PVR(Personal Video Recorder)이라는 저장 기능을 가진 TV를 이용해 양방향 서비스를 제공하고 있으며 원하는 키워드를 입력하면 해당 키워드가 나타나는 프로그램을 자동으로 녹화해주는 기능을 제공하고 있다.

7. 결론 및 향후연구

본 논문은 TV환경 하에서의 추천 시스템을 구현하고 실험을 통해 이의 효용성을 보였다. 충분한 데이터와 사용자를 확보할 수 있다면 좀 더 여러 가지 기능적인 요소들과 추천 기법을 추가할 수 있을 것이다. 추천 시스템 중 내용 기반 방식을 사용하는 시스템은 항상 추천의 대상이 되는 도메인에 여러 의존적인 특징들을 가지고 있다. 사용자의 성향 분석과 충분한 프로그램 정보와 같은 요인들은 시스템을 좀 더 정교하게 만드는데 많은 도움을 줄 수 있을 것이다. 또한 추천 시스템의 성능을 객관적으로 평가하기 위해서는 정보검색 분야와 같은 테스트 컬렉션이 필요하다. 현재 사용되고 있는 영화데이터 만으로 다른 도메인에 대한 추천기법을 평가하는 것이 곤란하므로 별도의 테스트 컬렉션이 마련되어야 할 것이다.

TV에는 TV만이 가진 여러 가지 특징이 있다. 프로그램의 스케줄이 일정 주기로 반복되므로 이를 추천에 이용할 수 있을 것이고 실제 시청자의 시청기록을 얻을 수 있으면 훨씬 더 나은 결과를 얻을 수 있을 것이다. 또한 협업 기반 방식이 가진 장점을 살리기 위해 연령별, 성별, 직업별 시청률 데이터 같은 것 등을 활용하여 추천 리스트에 포함시킨다면 내용 기반 방식이 가진 단점을 상당 부분 메워나갈 수 있을 것이다.

참고 문헌

[1] Paul Resnick and Hal R. Varian, Recommender Systems, CACM 40(3) pp.56-58, 1997.
 [2] Peter W. Foltz and Susan T. Dumais, Personalized Information Delivery: An Analysis of Information Filtering Methods, CACM 35(12) pp.51-60, 1992.
 [3] Resnick, P., Iacovou, N., Sushak, M., Bergstrom, P., and Riedl, J. GroupLens: An open architecture for collaborative filtering of netnews. Proceedings of the 1994 Computer Supported Collaborative Work Conference. 1994.
 [4] Marko Balabnovic and Yova Shoham, Fab: Content-Based, Collaborative Recommendation, CACM 40(3) pp.66-72, 1997.
 [5] Ricardo Baeza-Yates and Berthier Ribeiro-Neto, Modern Information Retrieval, Addison-Wesley, 1999.
 [6] <http://www.tv-anytime.org/>
 [7] R. Korfhage. Information Storage and Retrieval. John Wiley, New York, 1997.
 [8] Herlocker, J. Understanding and Improving Automated Collaborative Filtering Systems. Ph.D. Thesis, Computer Science Dept., University of Minnesota, 2000.
 [9] J. J. Rocchio Relevance feedback in information

- retrieval. In The Smart System-Experiments in Automatic Document Processing, pp.313-323, Prentice Hall, 1971.
- [10] Breese, Heckerman, and Kadie. Emperical analysis of predictive algorithms for collaborative filtering. Technical report, Microsoft Research, October 1998.
- [11] P. Melville and R. Mooney and R. Nagarajan, Content-boosted collaborative filtering, SIGIR 2001.
- [12] Barry Smyth and Paul Cotter, Surfing the Digital Wave: Generating Personalised {TV} Listings Using Collaborative, Case-Based Recommendation, LNCS 1650, 1999.
- [13] K. Kurapati and S. Gutta and D. Schaffer and J. Martino and J. Zimmerman, A multi-agent TV recommender, User Modeling, 2001.
- [14] Workshop on Personalization in Future TV 2001/2002.
- [15] <http://www.canalplus-technologies.com>
- [16] <http://www.tivo.com>

유 상 원

정보과학회논문지 : 컴퓨팅의 실제
제 9 권 제 5 호 참조



이 홍 래

2002년 3월~현재 서울대학교 전기.컴퓨터공학부 대학원 석사과정

이 형 동

정보과학회논문지 : 컴퓨팅의 실제
제 9 권 제 5 호 참조

김 형 주

정보과학회논문지 : 컴퓨팅의 실제
제 9 권 제 3 호 참조