

# IPTV환경에서 메뉴 선호 정보를 활용한 효율적인 VOD 추천 방법

오종수<sup>○</sup> 이성진 이수원

숭실대학교 컴퓨터학과

{polarojs<sup>○</sup>, ptnrev93}@mining.ssu.ac.kr, swlee@ssu.ac.kr

## An Efficient Method for Recommending VOD Contents by Using Menu Preferences in IPTV Environments

Jongsoo Oh<sup>○</sup>, Seongjin Lee, Soowon Lee

Dept of Computer Science, Soong-Sil University

### 1. 서 론

대량의 콘텐츠가 제공되는 환경에서 추천은 사용자의 콘텐츠 이용 편이성과 콘텐츠 소비 촉진이라는 점에서 필수 기능으로 요구되어지고 있다. 최근 급속도로 활성화 되고 있는 IPTV환경에서도 콘텐츠가 더욱 풍부해지면서 사용자의 선호에 맞는 콘텐츠를 선별하여 제공함으로써 선택에 도움을 줄 수 있는 추천의 필요성이 증대되고 있다. 본 논문에서는 기존의 다양한 연구에서의 콘텐츠 추천 방법을 IPTV환경에 적용하는데 따르는 확장성 문제[1]를 개선하기 위해 IPTV의 VOD 콘텐츠 이용 특징을 반영한 효율적인 추천 방법을 제안한다. 제안 방법은 사용자의 콘텐츠 접근 정보인 메뉴 정보와 사용이력 정보를 이용하여 개인의 선호를 추출하여 메뉴 선호에 기반한 콘텐츠 추천을 제공하며, IPTV의 VOD 콘텐츠 이용 특징인 시리즈 콘텐츠의 이용 특징을 반영하여 확장성 문제를 개선한다. 또한 개인 선호정보 기반 추천이 갖는 과거에 접했던 아이템과 다른, 새로운 아이템을 접할 기회가 제한되는 선호정보의 과잉 특성화(Over-specialization)문제[2]를 개선하기 위해 선호메뉴의 연관메뉴를 이용한 추천을 제안한다.

### 2. 본 론

본 연구에서는 사용자의 콘텐츠 이용 후 별도의 작업을 필요로 하는 직접적인 평가정보가 아닌 콘텐츠 시청 히스토리 정보와 사용자가 콘텐츠를 이용하기 위해 탐색하는 메뉴정보를 이용하여 사용자의 선호를 예측한다. 메뉴 선호정보를 이용한 콘텐츠 추천방법은 시청 히스토리로부터 사용자의 선호메뉴를 찾고, 그 선호메뉴의 특성에 따라 다양한 콘텐츠 추천방법을 적용한다. 또한 선호메뉴 내에서 콘텐츠 추천이 이루어질 수 없는 경우에는 선호메뉴의 연관메뉴를 이용한 콘텐츠 추천을 제공한다.

제안 시스템은 추천엔진과 학습엔진으로 구분된다. 추천엔진에서는 시청 히스토리과 메뉴 정보를 이용하여 메뉴 선호정보 학습, 인기 콘텐츠 추출, 연관메뉴 추출이 각각 수행되어 DB에 저장된다. 추천엔진에서는 학습엔진의 각 결과물과 콘텐츠 메타데이터를 이용하여 추천 요청에 따라 추천 리스트를 제공한다.

#### 2.1 제안 학습 방법

메뉴의 선호 정보학습은 메뉴 내의 콘텐츠 이용 빈도정보를 기반으로 이루어진다. 이러한 단순 빈도정보 기반의 메뉴 선호는 콘텐츠의 시리즈물 ‘몰아보기’의 특성[3]과 최근 선호 정보의 중요성을 반영하지 못한다. 이러한 문제점을 개선하기 위하여 본 연구에서는 선호 감소율을 적용한 메뉴 선호정보 학습 방법을 제안한다. 선호 감소율을 적용은 시청 히스토리를 일정 기간으로 나누어 기본 선호도를 학습하고, 기간별 기본 선호도에 선호 감소율을 식(1)과 같이 가중치로 곱하는 과정을 거쳐 반영된다. 최근성이 반영된 전체 기간의 선호도는 각각의 해당 기간 기본 선호도에 선호감소율을 가중치 곱으로 적용한 모든 기간의 선호도의 합으로 식(2)와 같이 계산된다.

$Preference_t = Preference \times W^t, (t \geq 0)$	(1)	$t \in T$ : 지난시간(기간) $W$ : 선호감소율( $0 < W \leq 1$ ) $Preference_t$ : $t$ 기간의 선호도 $Preference_{all}$ : 전체기간의 선호도
$Preference_{all} = \sum_{t \in T} Preference_t$	(2)	

인기 콘텐츠 추출은 전체 인기 콘텐츠와 메뉴 내의 인기 콘텐츠를 이용 빈도가 높은 순으로 순위화하여 DB에 저장하는 단계이다. 추출된 메뉴별 인기 콘텐츠는 단편 콘텐츠 추천의 가장 기본적인 인기 콘텐츠 추천에 활용되며, 전체 인기 콘텐츠는 선호메뉴 및 연관메뉴에서 추천이 불가능한 예외적인 상황에서의 인기 콘텐츠 추천에 활용된다.

연관메뉴 추출은 사용자가 시청을 시작한 시점부터 연속적으로 콘텐츠를 이용한 후 시청을 종료한 시점까지를 하나의 시청 단위로 보고 이를 시청 블럭으로 정의하여 시청 블럭 내의 메뉴의 이용 연관성을 추출하는 단계이다. 이를 위해 시청 블럭 내에서 같이 이용한 메뉴를 트랜잭션으로 추출하여 메뉴간 연관값(support, confidence, lift)을 계산하고 임계값 조절을 통해 필터링 과정을 거쳐 결과를 저장한다.

2.2 제안 추천 방법

제안한 추천 방법은 사용자의 추천 요청이 들어오면 선호메뉴를 선정하고 선호메뉴 내에서 선호할 만한 콘텐츠를 찾아서 제공하는 단계로 나눌 수 있다. 선호메뉴의 선정은 학습된 메뉴 선호정보를 이용하여 이루어진다. 선정된 선호메뉴는 시리즈 콘텐츠 편성 메뉴와 단편 콘텐츠 편성 메뉴로 구별되고 그에 따른 추천 방법을 적용한다. 선호메뉴에서 추천 콘텐츠를 선정하여 생성한 추천 리스트가 추천 요청 갯수에 미치지 못할 경우에 선호메뉴의 연관메뉴가 추출되며, 추출된 연관메뉴의 특성에 따른 콘텐츠 추천이 이루어진다.

본 논문에서는 선호메뉴별로 하나의 콘텐츠를 추천하는 것으로 추천 리스트 작성 방법을 제한한다. 그 이유는 한번에 추천을 제공하는 콘텐츠 수가 제한되어 있는 상황에서 시리즈 콘텐츠의 경우 한 시리즈에서 두 편 이상을 추천하는 것은 사용자의 다양한 콘텐츠 선택의 폭을 좁히게 되므로 좋은 추천이 아닐 것이기 때문이다. 선호메뉴가 시리즈 콘텐츠로 이루어진 메뉴인 경우에는 시리즈 규칙기반 추천을 적용하며 단편 콘텐츠로 이루어진 메뉴인 경우 다양한 추천 기법을 적용 할 수 있으나 여기서는 단편 콘텐츠에 대한 추천 성능 향상은 향후 연구로 남겨두고 있기 때문에 단순한 방법인 메뉴 내 인기 콘텐츠 추천을 적용한다.

선호메뉴에서 추천이 제한되는 경우 연관메뉴를 사용하여 추천을 수행한다. 선호메뉴에서 추천을 못하는 경우는 선호메뉴의 수가 추천 요청 수보다 부족한 경우 또는 사용자가 선호메뉴 내의 모든 콘텐츠를 소비한 경우이다. 선호메뉴에 대한 연관메뉴 선정은 식(3)의 연관메뉴의 선호도 예측값의 계산 및 순위화 과정을 거쳐 결정된다.

$$PredictedPreference(u_i, am_j) = \sum_{PM(u_i)} MenuPreference(u_i, pm_k) \times AssociationValue(pm_k, am_j) \quad (3)$$

*PredictedPreference*( $u_i, am_j$ ): 사용자  $u_i$ 의 연관메뉴  $am_j$ 의 선호도 예측값  
*MenuPreference*( $u_i, pm_k$ ): 사용자  $u_i$ 의 선호메뉴  $pm_k$ 의 메뉴 선호도  
*AssociationValue*( $pm_k, am_j$ ): 메뉴  $pm_k$ 와 메뉴  $am_j$ 의 연관값

3. 실험 및 결과

본 논문에서는 현재 서비스 중인 IPTV에서 2009년 12월의 실제 VOD 콘텐츠 제공 서비스의 운영 데이터를 이용하여 실험을 진행하였다. 이 중 3주의 데이터는 학습, 1주의 데이터는 평가에 활용하였다.

비교 평가를 위한 척도는 추천 횟수 중 추천 성공(추천 콘텐츠 중 하나 이상 평가 기간의 콘텐츠와 일치)의 비율을 나타내는 성공률(success rate)과 일반적으로 추천 정확도를 측정하는 precision, recall, f-measure를 이용하였다.

구분	방법	메뉴 선호			척도			
		메뉴 선호	시리즈 규칙	인기 콘텐츠	Success Rate	precision	recall	f-measure
1	전체 인기 콘텐츠	x	x	o	13.2%	0.0364	0.0048	0.0085
2	선호메뉴의 인기 콘텐츠 (시리즈 규칙X)	빈도	x	o	36.2%	0.0920	0.0141	0.0245
3	시리즈 규칙	x	o	o	55.3%	0.1904	0.0327	0.0558
4	선호메뉴(빈도)의 시리즈 규칙, 인기 콘텐츠	빈도	o	o	61.6%	0.1828	0.0307	0.0526
5	선호메뉴(감소율)의 시리즈 규칙, 인기 콘텐츠	감소율 0.1	o	o	65.2%	0.2014	0.0338	0.0579

[표 1] 메뉴 선호 기반 추천 성능 비교

척도	방법	추천수			
		5개	20개	60개	100개
Success Rate	인기콘텐츠+	65.3%	78.8%	81.2%	84.7%
	연관메뉴+	65.2%	78.8%	85.4%	87.2%
precision	인기콘텐츠+	0.2014	0.0929	0.0455	0.0345
	연관메뉴+	0.2014	0.0930	0.0483	0.0386
recall	인기콘텐츠+	0.0338	0.0572	0.0769	0.0973
	연관메뉴+	0.0338	0.0583	0.0860	0.1117
f-measure	인기콘텐츠+	0.0579	0.0708	0.0572	0.0509
	연관메뉴+	0.0579	0.0717	0.0619	0.0574

[표 2] 연관메뉴 추천 성능 비교

실험은 추천 콘텐츠 수를 5개로 기준을 설정하여 진행하였다. 예외적으로, 선호메뉴에서 추천이 이루어 지지 않는 상황을 가정한 연관메뉴 추천의 효과 검증에 위하여 추천 콘텐츠 수를 늘려가며 실험을 진행하였다. 메뉴 선호 기반 추천의 성능 비교를 통해 메뉴 선호를 이용한 추천 방법이 그렇지 않은 경우에 비해 성공률 면에서 좋은 결과를 보였으며, 메뉴 선호 학습에 시간에 따른 선호 감소율을 적용함으로써 메뉴 선호를 이용하지 않는 최근 시청의 시리즈 규칙보다 모든 평가 척도에서 더 좋은 성능을 보였다[표 1]. 또한 선호 메뉴에서 추천이 제한되는 경우를 가정한 실험에서 선호메뉴의 연관메뉴를 적용한 추천이 예외적인 상황에서 인기콘텐츠 추천을 제공하는 것보다 더 효과적인 추천이 된다는 것을 확인할 수 있었다[표 2].

향후 연구로는 제안한 방법의 최적화 실험과 단편 콘텐츠 추천 성능 개선을 위하여 아이템기반 협업필터링 등의 다양한 추천 방법 적용이 필요하며, 또한 영화, 음악 등 다양한 도메인의 콘텐츠 추천시스템에서 가장 많이 활용되고 있는 아이템기반 협업필터링 방법과 제안한 방법의 추천 성능을 비교 평가할 예정이다.

Acknowledge

본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2009년도 문화콘텐츠산업기술지원사업의 연구결과로 수행되었음.

참고문헌

[1] K. Yu, A. Schwaighofer, V. Tresp, X. Xu and H. Kriegel, Probabilistic Memory-Based Collaborative Filtering, IEEE Transactions on Knowledge and Data Engineering, vol.16 no.1, 2004. 1.  
 [2] B. Sheth, P. Maes, "Evolving Agents for Personalized Information Filtering", Proceedings of the 9th IEEE Conference on Artificial Intelligence for Applications, 1993.  
 [3] Lee S.W, Kim C.W, Determinants of IPTV-VOD Services Adoption, Korean Association for Communication & Information Studies, vol.46, pp.9-36, 2009. 5.