

맞춤형 방송의 통계적인 성능평가 방법

신사임*, 이종설*, 임태범*, 이석필*

*전자부품연구원

e-mail : {mirror, leejs, tlim, lspbio}@keti.re.kr

The evaluation mechanism for the personalized broadcasting services based on the evaluation measures in information retrieval

Saim Shin*, Jong-Soel Lee, Tae-Boem Lim*, Soek-Pil Lee*

*Korea Electronic Technology Institute

요 약

맞춤형 방송 솔루션은 디지털 TV 서비스 외에도, 향후 유망기술로 주목 받고 있는 IPTV, DMB 등의 다양한 방송 및 멀티미디어 서비스에 적용이 가능하다. 본 연구에서는 맞춤형 방송 서비스의 객관적인 평가를 위한 성능평가 방법을 제안한다. 정보검색 시스템 평가에 사용하는 정확률, 재현률과 역순위 평균 수치를 적용하여 맞춤형방송의 추천결과와 실제 시청자의 시청 프로그램의 차이를 분석하여 맞춤형 방송 시스템의 정확도와 사용자 만족도를 통계적으로 평가 가능한 메카니즘을 제안한다.

그 동안 평가가 이루어지지 않았던 맞춤형 방송 서비스를 복합적으로 평가하는 방법론을 제안함으로써, 맞춤형 방송 시스템의 지속적인 성능향상과 연구개발에 기여할 것이다. 또한, 맞춤형 방송 서비스의 산업화와 다양한 장비로의 확산에도 기여할 것으로 기대된다.

1. 서론

디지털 TV 환경이 대중화 되면서 메타데이터를 이용한 사용자 맞춤형 방송 서비스의 개발이 두드러지고 있다. 맞춤형방송은 사용자 기호, 단말 성능, 망 특성, 자연환경 - 시간, 장소, 사용자 기분상태 등 - 등을 포함한 사용환경 (usage environment)에 적합한 맞춤형방송 콘텐츠를 제공하고자 하는 서비스이다. 시청자의 측면에서 제한된 채널 선택만으로 방송 프로그램을 수신하는 수동적인 시청에서 벗어나 방송 서비스에 직접 참여하거나 또는 원하는 시간에 원하는 프로그램을 수신, 시청하고자 하는 욕구가 반영된 것이다.

높은 성능 및 사용자가 만족하는 맞춤형 방송 시스템의 개발을 위해 무엇보다 필요한 과정이 정확한 평가이다. 정확한 평가가 이루어져야 시스템의 부족한 부분을 찾아내어 보완하고, 다른 시스템과의 객관적인 성능 비교도 가능하다. 이러한 평가과정이 있어야 지속적인 시스템의 성능개선이 가능하다.

본 연구는 맞춤형 방송 시스템의 성능 측정을 위한 통계적이고 객관적인 평가 과정을 제안한다. 정확한 평가를 통하여 맞춤형방송 서비스의 개선이 이 연구의 목적이다.

2. 기존 연구

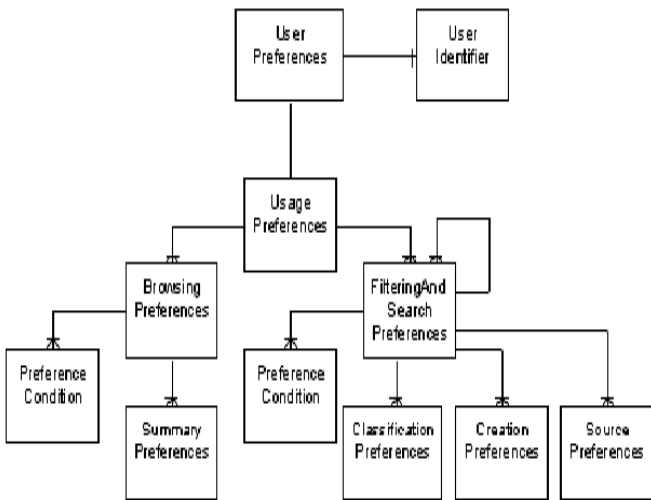
디지털 방송에서의 맞춤형 서비스란 사용자가 원하는 방송 프로그램만을 사용자가 원하는 시간에 볼 수

있게 해 두는 서비스를 말한다. 맞춤형 방송 시스템은 자동 사용자 프로파일링 (User profiling) 기술을 통하여 사용자의 시청패턴을 자동으로 학습하여 개인별 방송의 시청 환경을 제공해 준다.

2.1 TV-Anytime 의 맞춤형 서비스

맞춤형 방송 서비스의 규격을 통일하여 규정하고 있는 TV-Anytime 은 개개인의 사용자 선호정보를 바탕으로 한 Targeting 서비스에 대한 표준을 제정하였다. Targeting 서비스는 사용자 선호 정보에 추가적인 세부 정보 - 기분, 기후, 개인 소비정보 등 - 를 바탕으로 개개인의 현재 상황에 적합한 콘텐츠를 제공하는 것을 목적으로 한다. 이와 같은 서비스를 통해 TV-프로그램, 광고, 정보, 위치 등에 대한 targeting 서비스가 가능하다. TV-Anytime 에서 제공하는 맞춤형 방송 서비스를 위한 사용자 관련 메타데이터 표준으로는 사용자 히스토리 (Usage History)와 사용자 선호 정보 (User Preferences)가 있다. 그림 1은 사용자 히스토리 구성도이다. 이 메타데이터에는 사용자가 시청 과정에서 수행한 명령 및 대상 프로그램, 시간 등을 기록하여 사용자 별로 저장할 수 있는 자료구조를 제안하고 있다.

그림 1의 사용자 선호정보는 사용자 별로 선호 프로그램 추천에 유용한 정보들을 저장하고 있다. 선호하는 프로그램의 제목, 선호하는 프로그램의 주제, 장르, 선호채널, 선호시청 시간대 등을 기록하여 저장할 수 있도록 고안하였다.



(그림 1) TV-Anytime 의 사용자 선호정보 구조

2.2 맞춤형 방송의 성능 평가

맞춤형 방송 시스템은 지속적으로 개발되어 오고 연구가 진행되고 있지만, 성능평가를 위한 통계적인 measure 의 고안이나 방법론에 대한 고찰은 거의 이루어지지 않고 있다.

[1]은 인공지능의 학습개념을 도입하여 맞춤형 방송 서비스의 사용자 프로파일링 기술을 자동화하였다. Bayesian network 학습 알고리즘을 맞춤형 방송의 사용자 선호정보 자동 추출에 적용하여 서비스를 개발하였는데, 이의 평가를 위하여 기계학습 방법의 일반적인 평가방법인 결과의 에러율 (Error rate)을 사용하여 성능을 평가하였고 그 계산과정은 다음과 같다:

$$Errors(s) = \sum |\hat{\theta}_{x_{ij}} - \theta_{x_{ij}}| \text{----- (1)}$$

Im provement (%) =

$$\frac{Accuracy\ of\ nontraining - Accuracy\ of\ training}{Accuracy\ of\ nontraining} \times 100 \text{----- (2)}$$

식 (1)은 시스템이 추출한 사용자 선호정보의 각 아이템 값 - 예를 들어, 사용자 선호 장르의 드라마, 스포츠 등 -이 실제 사용자가 시청한 프로그램의 아이템 값과 일치할 조건부 확률을 계산하여 이를 에러율로 정의하여 자동 학습한 사용자 선호 정보값의 총 에러율을 계산하였다. 이 과정을 통하여 에러율이 낮거나 낮아지면 맞춤형 방송 시스템의 성능을 개선하였다고 평가한다.

식 (2)는 자동학습 부분을 배제한 기존 프로그램 추천 방식의 추천 결과 - [1]의 실험결과는 추천이 없이 임의로 제시 - 와 자동학습을 통한 추천목록의 각각의 에러율의 차이를 백분율화 하여서 맞춤형방송

시스템의 성능 향상 정도를 표현하고자 시도하였다.

2.3 정보검색 시스템의 성능 평가

역순위 평균 (Mean Reciprocal Rank)는 정보검색 시스템에서 정답으로 자동 추출한 웹 문서들의 적합성 정도를 평가하는 통계적인 평가 방법의 하나로, 웹 기반 정보검색 시스템의 성능 평가에 가장 일반적이고 표준이 되는 방법이다. 역순위 평균은 검색 시스템에 요청한 질의의 실제 정답을 포함하는 문서가 시스템이 제시한 검색 결과에서 몇 위에 위치하는지를 파악한다. 각 질의의 정답 순위의 역수를 합하여 그 평균을 계산하여, 제안하는 시스템의 정확도를 평가하는 방법이다. 웹 기반 정보검색 시스템은 사용자가 입력한 질의를 기반으로 사용자가 실제로 원하는 웹 콘텐츠들을 얼마나 정확하고 높은 순위에 추천하는 정도를 평가하기 위해 고안된 방법이다. 이 과정에서 정보검색 시스템은 사용자가 입력한 질의의 부족한 내용을 보완하고 사용자의 의도를 자동으로 파악하여 사용자가 원하는 문서를 빨리 찾을 수 있도록 적절한 순위화를 통하여 콘텐츠를 추천한다. MRR 값은 정보 검색 시스템의 질의보완, 매칭 정확도, 사용자 만족도 등을 복합적으로 평가하여 하나의 수치로 형상화할 수 있어 정보검색 시스템 평가에 표준화된 measure 로 자리잡게 되었다.

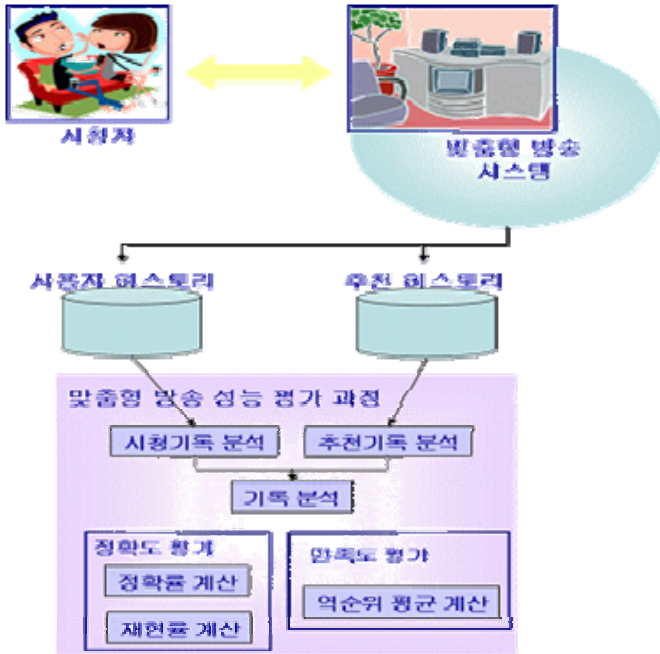
정확률 (Precision)과 재현률 (Recall)은 정보 검색뿐 아니라 인공지능의 자동학습, 자동분류 등 다양한 시스템의 평가 measure 로 일반적으로 사용하는 값이다. 정확률은 시스템의 정확도를, 재현률은 시스템의 coverage 를 표현하는 값이기 때문에, 시스템의 성능평가 시 두 값은 함께 사용하는 것이 바람직하다.

3. 맞춤형 방송의 평가요소 및 방법

본 연구에서는 맞춤형 방송 시스템의 객관적인 평가가 가능한 통계적인 평가 방법을 제안한다. 맞춤형 방송 시스템은 맞춤형 방송 관련 다양한 메타데이터 - MPEG, TV-Anytime -을 기반으로 사용자 선호정보를 자동으로 추출하여 사용자가 선호할 만한 프로그램과 환경을 사용자 별로 제공함으로써, 사용자들의 시청 만족도를 높이는 데 그 목적을 두고 있다. 이 과정에서 적합한 콘텐츠를 추출하여 사용자가 가장 좋아할 만한 콘텐츠를 순위화하여 제공하는 기능적인 측면으로 시스템을 분석한다면, 웹 기반 정보검색 시스템과 그 유형이 비슷하다고 할 수 있다.

맞춤형 방송 시스템의 성능평가를 위해, 사용자의 시청기록과 시스템의 추천 프로그램 기록을 분석하여 그 차이를 몇 가지 관점으로 분석하여 성능을 평가하였다. 또한, 맞춤형 방송의 추천 프로그램의 추출 및 순위화 과정이 사용자가 원하는 콘텐츠를 검색 및 순위화하여 제시하는 웹 기반 정보검색 시스템의 과정과 흡사하다. 그러므로, 정보검색 시스템의 성능평가 measure - 정확률, 재현률 및 역순위 평균값 -을 맞춤형 방송 평가에 적합하게 적용하여 사용하였다.

성능평가의 요소는 크게 두 가지 관점으로 나누어 서 맞춤형 방송의 추천 결과를 분석하여 평가하는 방 식을 제안한다. 시스템의 정확도 평가를 위한 분석과 사용자 만족도를 평가하기 위한 분석 과정으로 나누 어 제안하고 있다. 그림 2 는 제안하는 평가 방법의 과정을 설명하고 있다.



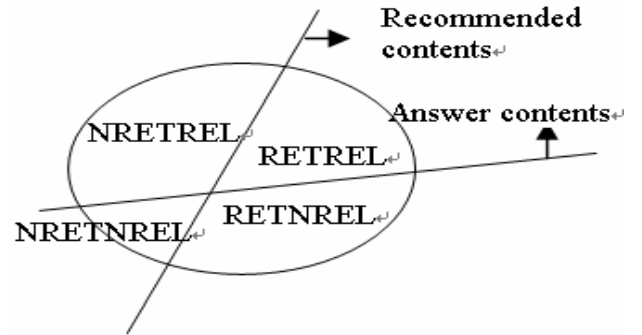
(그림 2) 맞춤형 방송 시스템의 성능평가 과정

성능 평가는 맞춤형 방송 시스템에서 각 시청자들 의 시청패턴을 분석하여 추출한 사용자 선호정보 (Use Preferences)를 기반으로 추천한 프로그램 히스토 리와 실제 시청자가 시청한 프로그램들을 사용자 히 스토리 (Usage history)로부터 추출 및 분석하여 이 결 과를 비교하는 과정으로 이루어진다. 평가 방식은 시 스템의 성능과 이에 따른 서비스가 얼마나 사용자의 기호에 맞게 이루어 지고 있는지의 사용자 만족도의 두 가지 관점에서 평가가 가능하도록 제안하고 있다. 따라서, 맞춤형 방송 시스템의 성능을 다양한 관점으 로 복합적으로 평가가 가능하며, 시스템의 특성 및 장단점을 정교하게 파악할 수 있을 것이다.

3.1 시스템 정확도 평가

제안하는 평가과정에서 맞춤형 방송 서비스의 정 확도 평가를 위하여 정확률과 재현률을 사용하도록 제 안한다. 맞춤형 방송 서비스의 성능 측정을 위해 적 용한 정확률은 시스템이 추천한 프로그램 리스트들 중 실제 사용자가 시청한 프로그램 목록과 일치하는 비율로 정의할 수 있다. 재현률은 사용자가 실제 시 청한 프로그램들 중 시스템이 추천한 프로그램 목록 과 일치하는 비율을 의미한다. 식 (3)과 (4)는 앞서 정 의한 정확률 (Precision)과 재현률 (Recall)을 수식화하 여 표현하고 있다 [4]. 식 (3)과 (4)의 정확률과 재현 른은 시스템이 추천한 방송 콘텐츠와 시청자가 실제

선택하여 시청한 콘텐츠가 어느 정도 일치하는지를 표현하고 있다. 정확률의 경우에는 맞춤형 방송 시스 템이 얼마나 정확하게 시청자가 선호하는 콘텐츠를 예측하는지를 나타내고, 재현률은 시청자가 실제로 선호하는 콘텐츠를 맞춤형 방송 시스템이 빠뜨리지 않고 얼마나 많이 추천하고 있는지를 말해주는 평가 measure 가 된다. 그러므로, 이 두 식을 통하여 각 맞춤형 방송 시스템의 정확도와 coverage 를 정확히 평 가할 수 있다.



$$Precision = \frac{RETREL}{RETREL + NRETREL} \quad (3)$$

$$Recall = \frac{RETREL}{RETREL + RETNREL} \quad (4)$$

3.2 시스템 만족도 평가

시스템 정확도 평가는 시스템이 추천한 결과에 사 용자가 실제로 시청한 콘텐츠를 얼마나 포함하고 있 는지 정도로 성능을 평가한다면, 시스템 만족도 평가 는 시청자가 실제로 시청한 콘텐츠를 얼마나 높은 순 위로 추천하고 있는지를 평가한다. 사용자가 실제 시 청한 프로그램들을 높은 순위로 추천한다면, 낮은 순 위로 추천하는 경우보다 시청자가 더 쉽게 콘텐츠를 발견하여 시청할 수 있게 된다. 그러므로, 시스템의 추천 리스트에 실제 시청한 콘텐츠가 포함되어 있다 고 하여도 높은 순위에 포함되어 있는지 여부 또한 사용자 만족도에 중요한 요소이다. 본 연구에서는 맞춤형 방송 시스템의 만족도 평가를 위하여 웹 기반 정보검색 방법에서 사용되는 역순위 평균값을 활용하 였다.

역순위 평균값은 식 (5)와 같다.

$$MRR = \frac{1}{n} \sum_{i=1}^n \frac{1}{r_i} \quad (5)$$

식 (5)에서, i 는 사용자의 시청기록에 나타나는 사 용자가 시청한 각각의 콘텐츠를 의미한다. r_i 은 i 가 맞춤형 방송 시스템이 추천한 목록에서 나타나는 추 천 순위를 말한다. 만약 각 시간대에 시청자가 실제 로 시청한 프로그램을 시스템이 모두 1 순위로 추천 하였다면, 평가하는 맞춤형 방송 시스템의 MRR 값은

1 로 수렴한다. 반대로, 시청자가 시청한 프로그램이 맞춤형 방송 시스템의 추천 리스트에 등장하지 않는다면, 해당 시스템의 사용자 만족도 MRR 값은 0 으로 수렴하게 된다 [2, 5].

웹 기반 정보검색 시스템은 사용자가 입력한 질의를 기반으로 사용자가 실제로 원하는 웹 콘텐츠들을 얼마나 정확하고 높은 순위에 추천하는 정도를 평가하기 위해 고안된 방법이다. 이 과정에서 정보검색 시스템은 사용자가 입력한 질의의 부족한 내용을 보완하고 사용자의 의도를 자동으로 파악하여 사용자가 원하는 문서를 빨리 찾을 수 있도록 적절한 순위화를 통하여 콘텐츠를 추천한다. 맞춤형 방송의 콘텐츠 추천 과정은 사용자의 선호정보를 기반으로 사용자의 선호정보에 적합한 콘텐츠를 수많은 방송 콘텐츠들 중에 검색하여 제공한다고도 할 수 있다. 이 과정에서 자동 학습을 통하여 부족한 사용자 선호정보를 학습하여 질의를 보완하는 등 정보검색 시스템과 비슷한 역할을 한다고 말할 수 있다. 즉, 정보검색 시스템의 다양한 기능을 하나의 수치화하여 평가하고 있는 역순위 평균값은 맞춤형 방송의 성능 평가에도 효율적으로 적용할 수 있다.

4. 결론

맞춤형 방송은 사용자와 밀접하게 상호작용하는 HCI (Human Computer Interface) 시스템 중 하나이다. 그러나, 다양한 기능을 제공하고 그 만족도 또한 사용자에게 따라 주관적이기 때문에 평가방법의 확립에 어려운 점이 있다. 서비스의 특성 상 시스템 평가가 어렵기 때문에, 개발된 시스템의 평가는 생략되거나 일회적이었다. 본 논문은 맞춤형 방송 시스템의 평가를 위한 방법론을 제안하고 있다.

사용자의 선호도 패턴을 수집하는 인공지능 에이전트의 성능뿐 아니라, 사용자의 시청 만족도 또한 객관적으로 평가할 수 있어야 바람직한 평가방법이라고 할 수 있다. 본 연구에서는 비슷하게 시스템의 성능 - 정확도 및 coverage - 와 사용자 만족도를 평가 요소로 측정하고 있는 정보검색 시스템의 평가 방식을 맞춤형 방송 시스템의 평가에 적합하게 적용하여 제안하였다.

맞춤형 방송 시스템의 통계적인 평가가 가능해 진다면, 다양한 시스템과 방법론의 객관적인 성능비교가 가능해 지기 때문에, 지속적인 기술개발이 가능하다. 또한, 상용화 과정의 시스템 성능 안정화에도 기여하게 될 것이다.

참고문헌

- [1] Sanggil Kang, Jeongyeon Lim, Munchurl Kim: Modeling the user preference of broadcasting content using Bayesian networks. Journal of electronic Imaging, Vol 14 (2). (2006)
- [2] Nick Craswell, David Hawking: Overview of the TREC-2004 Web Track. Workshop on TREC-13 (2003)
- [3] SP003v1.3 Part A: Metadata. TV-Anytime Forum (2002)
- [4] Cronbach L. J.: On the non-rational application of

information measures in psychology, in H Quastler, ed. Information Theory in Psychology: Problems and Methods. Free Press, Glencoe, Illinois, pp. 14- 30 (1954).

- [5] 신사임, 이종설, 이석필: 사용자 패턴의 자동추출을 통한 TV-Anytime 기반 사용자 선호정보 관리 시스템, 한국정보과학회 학술대회 (2006)
- [6] Saim Shin, Jong-Sul Lee, Tae-Beom Lim, Seok-Pil Lee: Automatic management of the user preferences with the information retrieval techniques in the personalized broadcasting services, ICUT (2007)