

협력적 여과 방식을 이용한 SCORM 기반 학습 콘텐츠 추천

현영순*, 조동섭
이화여자대학교 컴퓨터학과
e-mail:toyloveys@ewhain.net, dscho@ewha.ac.kr

SCORM based Learning Contents Recommendation using Collaborative Filtering

Young-Soon Hyun*, Dong-Sub Cho
Dept of Computer Science & Engineering, Ewha Womans
University

요 약

SCORM의 Content Repository는 Asset이나 콘텐츠의 Metadata를 가지고 콘텐츠나 Asset을 검색할 수 있도록 한다. 이런 Metadata 기반 검색은 아주 많은 콘텐츠를 대상으로 검색을 할 경우, 검색을 통한 콘텐츠 결과가 너무 많을 경우 결과 내에서 재검색을 하는데 많은 시간을 들일 수 있다는 단점이 있다. 본 논문에서는 검색 효율을 높이기 위한 방법으로 SCORM 기반 LMS에 협력 필터링 방법을 적용한 시스템을 제안하였다.

1. 서론

e-Learning 교육의 확산으로 인해 학습자 측면에서 시간과비용이 절감된 효과를 가져왔으나, 다양한 학습 콘텐츠가 요구되면서 콘텐츠를 생성하는데 많은 시간과 비용이 소요되었다. 또한 콘텐츠가 그것을 제작하는데 사용된 특정 LMS(Learning Management System)에 종속됨으로써 다른 LMS에 사용할 수 없으며, 다른 LMS에서 사용되던 콘텐츠 공유도 현실적으로 불가능했다.

이러한 LMS의 문제점을 해결하고자 ADL(Advanced Distributed Learning)에서는 SCORM(Sharable Content Object Reference Model)이라는 표준안을 내놓았다. SCORM의 궁극적인 목적은 콘텐츠와 시스템간의 상호운용성, 재사용성, 제어성 등을 확보함으로써 사용자에 대한 질 높은 서비스를 제공하는데 있다[1].

본 논문에서는 사용자가 SCORM기반 LMS 상의 수많은 교육용 콘텐츠를 검색할 때 원하는 콘텐츠에 쉽고 빠르게 접근할 수 있도록 보통의 협력 필터

링 방법을 이용하되 협력 필터링을 적용할 수 없는 콘텐츠에 대해서는 사용자가 가장 많이 본 콘텐츠를 추천하는 시스템을 제안하였다.

2절에서는 관련연구로 현재까지 연구된 SCORM과 협력 필터링 기법에 대해 살펴보고, 3절에서는 협력 필터링 방법을 적용한 SCORM 기반 학습 콘텐츠 추천 방법을 제안한다. 마지막으로 4절에서 본 논문에 대해 결론짓는다.

2. 관련연구

2.1 가상교육 시스템과 SCORM

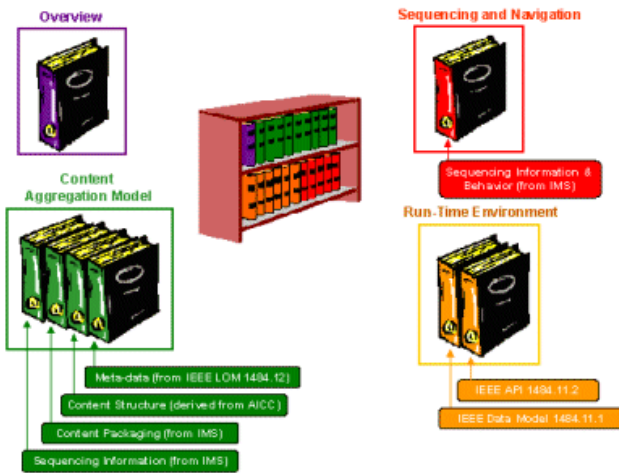
가상교육(Virtual Education)은 웹이 갖고 있는 특성과 자원의 활용을 통해 전자화된 가상공간에서 교수와 학습자가 참여하여 교수자, 학습자, 운영자 및 교육프로그램 간의 다양한 상호작용이 시공적인 제약 없이 수행되는 의도적이면서도 계획된 교수-학습 활동이라 할 수 있다.

웹은 다양한 형태의 상호작용을 가능하게 한다는 점에서 교수자와 학습자 또는 학습자와 학습자간의

역동적인 상호작용이 이루어질 수 있도록 지원할 수 있으며, 학습자들의 의사소통 기술뿐만 아니라 문제 해결 능력과 학습 능력을 개선할 수 있다는 점에서 많은 교육 단체에서 가상교육을 실행해 나가고 있다.

가상교육 시스템 표준으로는 각기 다른 단체에서 여러 표준안이 제안되고 있지만 SCORM은 정부, 기업, 교육기관들이 참여하여 각 표준안의 장점들을 모아 하나로 통합된 형태의 표준안이다. 이는 1997년 미 국방성에서 가상교육의 표준안으로 제안되었으며 이후 여러 업체들이 참여하여 현재까지 계속 발전해 나가고 있다. SCORM은 웹 공간에 존재하는 공유할 수 있는 콘텐츠들을 모아서 실시간으로 주문형(on-demand)으로 조합을 할 수 있도록 표준을 정해 다양하고 수준 높은 교육을 언제, 어디서나, 누구나 접하도록 하고자 하는 것이 목표이다.

현재 2004년 7월에 발표된 최신 버전인 SCORM version 1.3.1은 아래 그림 1과 같이 크게 Overview, Content Aggregation Model, Run-Time Environment, Sequencing and Navigation의 네 부분으로 이루어져 있다[2].

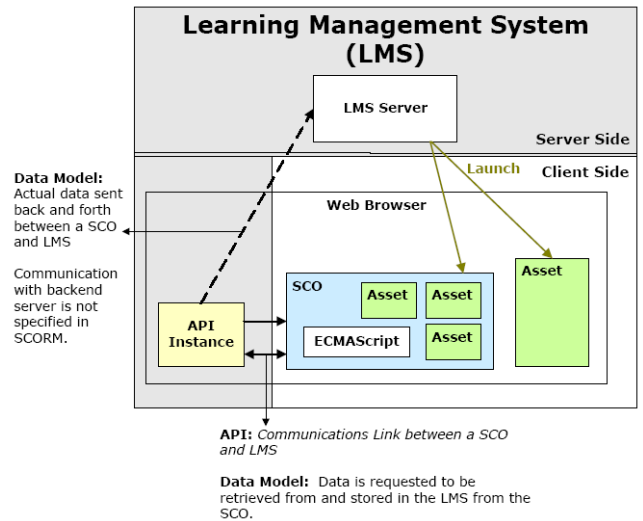


(그림 1) SCORM의 구성

2.2 SCORM Run-Time Environment

SCORM의 목적은 LMS에 관계없이 콘텐츠 객체들을 재사용 할 수 있게 하고 상호 운용이 가능하도록 하는 것이다. 이것을 가능하게 하기 위해, 콘텐츠를 내보내고 관리하는 공통적인 방법, LMS와 통신하기 위해 콘텐츠 객체들에 대한 공통 방법 그리고 통신의 기초를 형성하는 미리 정의된 언어나 어휘가 있어야만 한다. 그림 2에서 보여지는 것처럼, RTE

의 이 세 가지 양상은 Launch와 Application Programming interface (API) 그리고 Data Model이다.



(그림 2) 개념적인 SCORM 기반 실행환경

Launch 과정은 웹 기반 콘텐츠 객체들을 시작하기 위한 LMS들을 위해 공통된 방법을 규정한다. 콘텐츠 객체라는 용어는 학습자를 위하여 론치(launch)될 수 있는 콘텐츠의 어떤 부분을 기술하기 위해 일반적으로 사용된다. SCORM에서, 콘텐츠 객체들의 2가지 종류는 SCOs와 Assets이다. Launch 과정은 론치된 콘텐츠 객체와 LMS사이의 통신의 설립에 대한 절차와 책임을 규정한다. 통신 메커니즘은 공통 API의 사용을 통해 표준화된다.

API는 콘텐츠 객체와 LMS사이의 개념적인 통신 상태(예를 들면, 초기화, 종료, 에러 상황)를 LMS에게 알리는 통신 메커니즘이고, LMS와 SCO사이의 자료(예를 들면, 점수, 시간 제한, 등등)를 검색하고 저장하기 위해서 사용된다.

Data Model은 퀴즈 혹은 테스트와 같은 평가로부터 SCO의 완성 상태 또는 점수 같은 SCO를 위해 추적된 정보를 규정하도록 사용된 데이터 모델 엘리먼트들의 표준화된 집합이다[3].

2.3 협력 필터링(Collaborative Filtering)

다른 사람들의 관심 사항을 예측하기 위해 동료의 의견을 사용하는 기법으로, 제록스 팔로알토 연구소(PARC)의 Nichos등에 의해서 개발된 Tapestry가 협력 필터링을 적용한 최초의 시스템이다. 그리고 미네소타 대학의 GroupLens는 현재로서 가장 유명한 협력 필터링 시스템 형태이다. 협력 필터링 시

시스템은 같은 취향이나 취미를 가진 사람들의 정보를 이용해 추천을 할 때 도움을 줄 수 있는 가장 널리 사용되는 시스템이다. 링고 시스템은 음악 앨범 추천용으로, 무비렌즈 시스템은 영화 추천용으로, 제터 시스템은 조크 추천용으로, 그리고 플라이캐스팅은 온라인 라디오 추천용으로 개발된 협력 필터링 시스템들이다. 최근에는 인터넷 서점인 Amazon이나 인터넷 CD 상점인 CDNow, 그리고 인터넷 영화 추천 사이트인 MovieFinder 등에서 협력 필터링 방법을 적용시켜 성공을 거두고 있다. 이처럼 협력 필터링은 정보 필터링 응용분야나 상거래 응용분야에서 연구나 실용적 측면 모두에서 성공을 거두고 있으나 다음과 같은 근본적으로 해결해야 할 몇 가지 문제점을 갖고 있다.

- 초기 평가 문제(Early rate problem) : 순수한 협동 여과는 새로운 항목이 등장하여 예측에 기반이 되는 사용자들의 평가들이 없을 경우, 이 항목에 대한 예측을 제공할 수 없다. 더군다나, 특정 항목에 대한 초기 예측들은 종종 부정확하다. 왜냐하면 예측에 사용되는 평가들이 매우 적기 때문이다. 이와 유사하게, 심지어 이미 설치되어 운영되는 시스템조차도 시스템에 가입하는 모든 새로운 사용자들에게 불충분한 예측들을 제공할 것이다. 초기 평가 문제의 극단적인 예로서, 협력 필터링 시스템이 처음 시작될 때 모든 사용자는 모든 항목에 대해서 초기 평가 문제로부터 영향을 받는다.
- 희소성 문제(Sparsity problem) : 많은 정보 도메인에서는, 항목들의 개수는 개별 사용자들이 소화할 수 있는 개수를 훨씬 초과한다. 따라서, 모든 사용자에게 대한 모든 항목들의 평가들을 포함하고 있는 행렬들은 매우 드문드문한 분포성을 띤다. 이는 사용자간 유사도 계산 시 부정확성을 초래하며 결과적으로 협력 예측의 질을 떨어뜨리게 된다.
- 확장성(Scalability) : 협력 필터링 분야에서 주로 사용하는 최근접 이웃 알고리즘(Nearest neighbor Algorithm)은 사용자와 항목 수에 비례해서 계산 시간이 비례한다. 따라서, 사용자 수와 항목 수가 수백만이나 되는 환경 하에서는 이러한 계산 시간이 치명적일 수 있다[4].

이러한 문제점으로 인해 본 논문에서는 협력 필터링을 적용하기 어려운 경우에는 사용자가 가장 많이 본 콘텐츠를 추천하도록 하였다.

3. 협력 필터링을 이용한 콘텐츠 추천 시스템

ADL의 SCORM 표준안에서는 교육용 콘텐츠의 메타 데이터 즉 콘텐츠에 대해서 기술한 데이터의 형식에 관한 표준안을 포함하고 있다. 이 메타 데이터는 콘텐츠의 제목, 설명, 저자 등의 콘텐츠에 대한 내용을 XML 형식으로 표현하고 있고 이는 콘텐츠의 검색에 유용하게 활용될 수 있다. 하지만 수많은 콘텐츠들이 LMS 상에 있다면 메타데이터 기반 검색을 하더라도 너무 많은 콘텐츠가 검색될 수 있고 검색 결과 내에서 재 검색 하는데 많은 시간을 들이게 될 수 있다. 따라서 본 논문에서는 메타데이터를 이용한 콘텐츠 검색을 하되 현재 사용자와 유사한 성향을 가지는 이웃 사용자에게 의해 높은 선호도를 보인 콘텐츠를 추천한다.

본 논문에서 제안하는 협력 필터링을 이용한 콘텐츠 검색 시스템 구성은 다음과 같다.

3.1 사용자 선호도 저장

웹 서버는 사용자로부터 콘텐츠에 대한 선호도를 입력 받아서 서버의 데이터베이스에 저장한다. 사용자가 한번 선호도를 표시한 것은 다시 선호도를 묻지 않도록 하여 데이터의 중복을 피하도록 설계하였다.

3.2 유사도 계산

현재 사용자와 유사한 히스토리 정보를 가지는 이웃들의 집합을 찾기 위해 통계적 방법을 사용한다. 유사한 이웃들이 찾아지면 특정 콘텐츠를 현재 사용자에게 추천해주기 위하여 여러 알고리즘들을 사용하게 된다.

3.2.1 피어슨 상관관계 공식을 이용한 유사도

가장 많이 사용되는 가중치 측정은 피어슨 상관관계 방법이다. 피어슨 상관관계 방법은 두 변수간에 존재하는 선형 관계의 정도를 측정한다. 피어슨 상관관계수는 선형 리그레션 모델에서 유도되며 상관관계는 선형적이며 예러는 서로 독립적이고 평균 0인 확률 분포와 독립 변수의 모든 설정에 대한 상수 편차를 가진다는 가정에 기반을 둔다.

$$w_{a,k} = \frac{\sum_i (r_{a,j} - \bar{r}_a)(r_{k,j} - \bar{r}_k)}{\sqrt{\sum_i (r_{a,j} - \bar{r}_a)^2 * \sum_i (r_{k,j} - \bar{r}_k)^2}} \quad \dots \text{식(1)}$$

여기서 j는 사용자 a와 사용자 k가 모두 선호를 매긴 정보를 의미하고 $r_{x,y}$ 는 사용자 x의 정보 y에

대한 선호도를 나타내며 \bar{r}_x 는 사용자 x의 전체 정보에 대한 평균 선호도이다.

3.3 협력 예측

사용자의 항목에 대한 예측을 구하기 위해 GroupLens에서 제안한 식을 사용하였다. 여기서는, 항목에 대한 예측은 이웃의 평균값으로부터 편차의 가중치 평균을 수행함으로써 계산된다.

$$P_{aj} = \bar{v}_a + \frac{\sum_{i=1}^n w_{a,i} (v_{i,j} - \bar{v}_i)}{\sum_{i=1}^n w_{a,i}} \quad \text{..식(2)}$$

여기서 j는 사용자 a와 사용자 k가 모두 선호를 매긴 정보를 의미하고 $r_{x,y}$ 는 사용자 x의 정보 y에 대한 선호도를 나타내며 \bar{r}_x 는 사용자 x의 전체 정보에 대한 평균 선호도이다.

3.4 사용자가 가장 많이 본 콘텐츠 추천

정해진 임계치 이상의 사용자 선호도 값이 주어지지 않았을 경우, 사용자가 가장 많이 본 콘텐츠를 추천한다. 이때 콘텐츠의 조회수는 한 사용자가 특정 콘텐츠에 한번 접속한 횟수를 의미한다.

4. 결론

본 논문에서는 사용자가 SCORM기반 LMS 상의 수많은 교육용 콘텐츠들을 검색할 때 원하는 콘텐츠에 쉽고 빠르게 접근할 수 있도록 협력 필터링 방법을 적용하였다. 그러나 정해진 임계치 이상의 사용자 선호도 값이 주어지지 않았을 경우, 사용자가 가장 많이 본 콘텐츠를 추천한다. 이를 통해 협력 필터링에서 발생하는 여러 가지 문제점을 해결할 수 있다.

참고문헌

- [1] 김용만, 김현철, "SCORM 스펙을 이용한 학습관리 시스템의 설계 및 구현", 한국 멀티미디어학회지, 제 6권, 제 4호, pp.23 - 32, 2002. 12.
- [2] ADL(2004), "Sharable Content object Reference Model Version 1.3 - "SCORM overview", Advanced Distributed Learning.
- [3] ADL(2004), "Sharable Content object Reference Model Version 1.3 - "SCORM Run-Time Environment", Advanced Distributed Learning.
- [4] 김병만, 이경, 김시관, 임은기, 김주연, "추천시스템을 위한 내용기반 필터링과 협력필터링의 새로운

결합 기법", 정보과학회논문지 소프트웨어 및 응용, 제 31권, 제 3호, 2004. 3.

[5] ADL(2004), "Sharable Content Aggregation Model Version 1.3.