

연관규칙을 응용한 SCORM 기반 학습 콘텐츠

현영순*, 조동섭
이화여자대학교 컴퓨터학과

SCORM Based Recommendation of Learning Contents using Association Rule Mining

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

Abstract - 본 논문에서는 SCORM을 기반으로 하는 LMS 상에 수많은 콘텐츠들이 있을 경우, 적은 노력으로도 원하는 콘텐츠에 접근할 수 있도록 도움을 주는 콘텐츠 추천 기법을 제안하였다. 이 기법은 각 학습자별로 콘텐츠 이용도 성향을 분석한 후 분석된 결과를 바탕으로 사용자에게 현재 이용하고 있는 콘텐츠와 가장 연관성이 높다고 판단되는 콘텐츠를 연관규칙을 응용한 방법을 이용하여 추천한다.

서 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 단계에서 여러 표준안이 제안되고 있지만 SCORM은 정부, 기업, 교육 기관들이 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.

Overview 부분에서는 역사와 SCORM의 목적에 대해서 Content Aggregation Model 부분에서는 교육용 컴포넌트와 그 컴포넌트를 여러 시스템에서 사용하기 위해 포장(package)하는 방법에 대해서 그리고 Sequencing and Navigation에서는 그 컴포넌트를 찾는 방법과 컴포넌트의 sequencing rule 등에 대해서 설명하고 있다.

1. 서 론

e-learning 교육의 확산으로 인해 학습자 측면에서 시간과 비용이 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 ADI(Advanced Distributed Learning)에서는 SCORM(Sharable Content object Reference Model)이라는 표준안을 내놓았다. SCORM의 궁극적인 목적은 콘텐츠와 시스템간의 상호운용성, 재사용성, 제어성 등을 확보함으로써 사용자에게 대한 질 높은 서비스를 제공하는데 있다[1].

본 논문에서는 사용자가 SCORM 기반 LMS 상의 수많은 교육용 콘텐츠들을 검색할 때 원하는 콘텐츠에 쉽고 빠르게 접근할 수 있도록 각 학습자별로 콘텐츠 이용도 성향을 분석하고, 그 학습자가 현재 이용하고 있는 콘텐츠와 연관성이 높은 콘텐츠들에 대해서 우선순위를 부여하여 검색결과에 나타내어 주는 추천 시스템을 제안하였다.

2절에서는 관련연구로 현재까지 연구된 SCORM과 연관규칙 기법에 대해 살펴보고, 연관규칙을 응용한 SCORM 기반 학습 콘텐츠 추천 방법을 제안한다. 마지막으로 3절에서는 결론과 향후 연구과제에 대해 언급한다.

2. 본 론

2.1 관련연구

2.1.1 가상교육 시스템과 SCORM

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

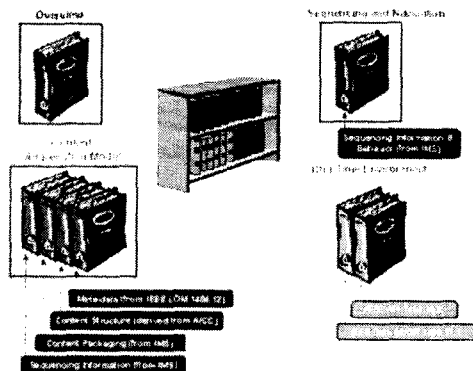


그림 1. SCORM의 구성

또 RTE 부분에서는 콘텐츠를 내보내는 프로세스, 콘텐츠와 LMS 사이의 통신, 학습자에 대한 정보를 보내기

이 논문은 2005년도 두뇌한국21사업에 의하여 지원되었음

위한 표준화된 데이터 모델 등과 같은 실행환경을 관리하기 위해 필요한 사항에 대해 나타나 있다. SCORM에서 사용되는 몇몇 용어들에 대한 정의는 다음과 같다.

Asset이란 미디어, 텍스트, 이미지, 사운드, 웹페이지, 평가 객체 그리고 웹 클라이언트에게 전달될 수 있는 데이터들의 전자적 표현이다.

SCO(Sharable Content Objects)는 독립적이고, 교수자료 부분으로 정의된 Asset의 집합이다. SCO는 LMS에 의하여 전달과 추적할 수 있는 교육의 가장 작은 논리적 단위이다. 현재 SCORM 파라미터 아래, SCO는 직접적으로 launch나 다른 SCO를 접근할 수 없다. 모든 inter-SCO sequencing은 LMS에 의해 지시된다.

Content Repository는 일반적으로 Asset이나 콘텐츠의 Metadata를 추가할 수 있는 기능을 포함하고 있으며 Metadata를 가지고 콘텐츠나 Asset을 검색할 수 있도록 한다. Content Repository는 의도적으로 재양이나 사건에서 데이터를 보호하고 대용량 저장장치의 비용을 줄이기 위해 분산되어 있다[2].

2.1.2 SCORM 기반 LMS

SCORM 기반 LMS란 SCORM의 규격에 맞는 LMS를 말한다. SCORM은 콘텐츠를 재사용 가능하게 하고 여러 개의 LMS 간의 상호 운용성을 제공하는 것을 목적으로 하는데 이를 가능하게 하기 위해서 콘텐츠를 내보내고 관리하는 공통적인 방법, 콘텐츠와 LMS 간의 통신을 위한 공통된 방법 그리고 LMS와 콘텐츠 사이에 상호 교환될 수 있는 미리 정의된 데이터 요소(data element)가 있어야만 한다. SCORM 실행 환경에서의 이 세가지 요소가 바로 그림 2에서 볼 수 있는 Launch, Application Programming Interface(API) 그리고 Data Model이다[3].

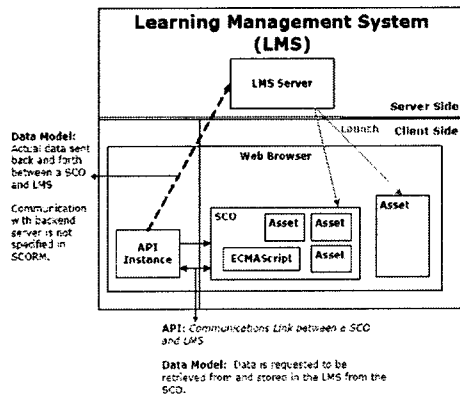


그림 2. 개념적인 SCORM 기반 LMS 실행환경

Launch 메커니즘은 LMS가 웹 기반의 학습자원을 시작하기 위한 공통의 방법을 정의한다. 통신 프로토콜은 공통 API의 사용을 통해 표준화된다.

API는 LMS와 SCO 사이에서 데이터(e.g., 점수, 시간, 제한 등)를 가져오고 저장하기 위해 사용되며 학습자원의 상태(e.g., 초기화, 종료, 예러 상황의 상태)를 LMS에 알리는 통신 메커니즘이다.

Data Model은 평가(e.g., quiz나 test)의 completion status 또는 score와 같은 SCO를 위한 정보를 정의하는데 사용된 표준화된 데이터 요소들의 집합이다.

이렇게 SCORM 기반 LMS는 SCORM에서 정하고 있는 Launch 메커니즘, API, Data Model의 규격을 따르는 LMS를 말하고 다음 그림3은 SCORM 기반 LMS의 예를 나타낸 것이다.

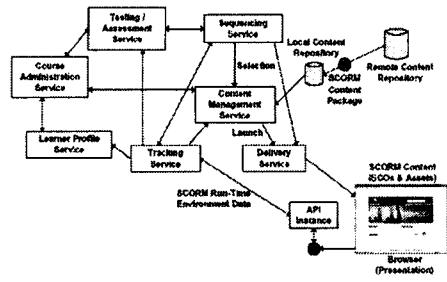


그림 3. 일반적인 LMS 모델

2.1.3 연관규칙 탐사

연관규칙은 빈발 항목집합(Large Itemset or Frequent Itemset)에 대한 연관규칙을 생성해내는 작업으로 나눌 수 있다. 전체 항목집합(상품들의 집합)이라 불리는 문자들의 집합 $I=(i_1, i_2, \dots, i_m)$ 이 주어지면, D는 트랜잭션(Transaction; 판매기록 - 판매데이터블의 한 레코드)의 집합이고, 각 트랜잭션 T는 T의 항목들의 집합이라고 정의된다. 한 트랜잭션에서 구매한 항목들의 양에는 상관하지 않고, 각 항목은 그 항목의 구매 여부만을 나타내는 이진변수임을 의미한다. 각 트랜잭션은 TID라 불리는 식별자를 지니고 있다. X를 항목들의 집합(Itemset)이라 하고 트랜잭션 T가 X를 포함한다면 $X \subseteq T$ 이라 한다. 연관규칙은 $X \Rightarrow Y$ 의 형식으로 표시되고, $X \subseteq I, Y \subseteq I$ 및 $X \cap Y = \emptyset$ 는 공집합이다. 또한 $X \Rightarrow Y$ 신뢰도(C Confidence) c를 가지고 있다는 것은 X를 포함하는 D의 트랜잭션들 중 c%가 Y도 포함하고 있음을 의미한다. 또한 $X \Rightarrow Y$ 는 트랜잭션 집합 D내에 지지도(Support) s%를 갖는데 이는 D의 트랜잭션들 중 s%는 X를 포함하고 있음을 뜻한다. 최소 지지도(Minimal Support) 이상을 갖는 항목집합을 빈발 항목집합이라 한다. K개의 항목들로 이루어진 빈발 항목집합을 빈발 k-항목집합이라고 한다. 빈발 k-항목집합들의 집합을 Lk라 하고 이를 위한 후보 k-항목집합들의 집합을 Ck라고 한다.

연관규칙을 발견하는 과정은 다음 두 단계로 구성된다. 첫째는 빈발항목 집합들을 찾는 것이다. 즉 주어진 최소 지지도 이상의 트랜잭션들의 지지를 갖는 항목 집합들의 모든 집합을 찾는 것이며, 둘째는 데이터베이스에 대한 연관규칙을 찾기 위해 빈발항목집합을 사용한다. 발견된 연관규칙은 지지도 및 신뢰도를 가지는 수치 값으로 정량화 된다[4].

2.2 연관규칙을 응용한 콘텐츠 추천

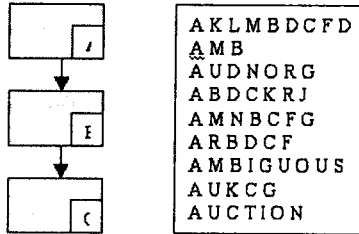
이 기법은 학습자 별로 콘텐츠 이용도 성향을 분석한 후 분석된 결과를 바탕으로 사용자에게 현재 이용하고 있는 콘텐츠와 가장 연관성이 높다고 판단되는 콘텐츠를 연관규칙을 응용한 방법을 이용하여 추천한다.

그림 4를 보면, 현재 사용자는 C 콘텐츠를 있고 이 콘텐츠에 오기 전에 B와 A를 역순으로 거쳤다. 초기 데이터베이스 DB는 기존 사용자들이 학습 콘텐츠 즉 하나 이상의 SCO로 이루어진 SCORM 콘텐츠를 이용한 경로 정보를 하나의 트랜잭션으로 묶어 모은 것이다. Count는 Next Content가 나온 횟수이다.

초기 데이터베이스 DB에서 현재 콘텐츠 C가 있는 항목들을 선택하여 데이터베이스 D₀를 구성한다. 선택된 각 트랜잭션 내에서 C의 Next Content를 찾아 Count를 계산한 P₀를 만든다. 다음 반복에서는 D₀를 스캔하여 현재 콘텐츠 C 이전의 학습한 콘텐츠 C의 두 번째 이전 콘텐츠인 A가 C의 이전에 있는 트랜잭션들을 선택하여 D₂를 구성한다. 선택된 각 트랜잭션 내에서 C의 Next Content를 찾아 Count를 계산한 P₂를 구성한다. 더 이상 탐색할 콘텐츠가 없으므로 반복을 중지한다. 탐색할 컨

텐츠가 더 있더라도 데이터베이스 D_k 가 더 이상 생성되지 않는다면 반복을 중지한다.

이렇게 생성된 P_k 들 중에서 가장 큰 k 값을 가진 $P(\max k)$ 에 있는 Next Content를 추천한다. Next Content가 여러 개일 경우에는 Count가 가장 큰 콘텐츠를 추천한다.



현재 사용자의 경로정보 기존 사용자의 경로정보
(Database DB)

그림 4. 사용자 경로정보와 초기 데이터베이스

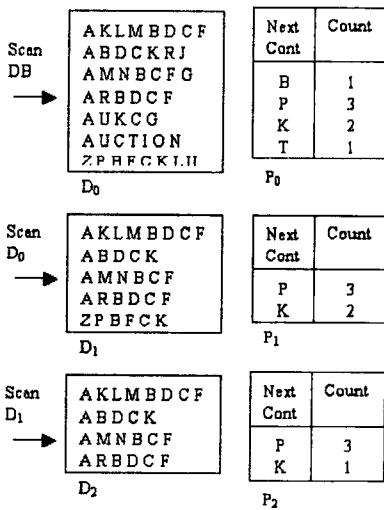


그림 5. 추천 콘텐츠 계산 과정

3. 결 론

본 논문에서는 SCORM 기반 LMS에서 사용자가 지나 온 경로정보에 따라 동적으로 관련된 콘텐츠를 계산하여 추천하는 기법을 설명하였다.

향후 연구로는 데이터베이스 DB를 전체 사용자의 경로정보로 보지 않고, 사용자를 그룹화하여 각 그룹의 경로 정보로 본다면 현재 사용자에게 좀 더 개인화 된 콘텐츠를 추천할 수 있을 것이다.

[참 고 문 헌]

- [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] 임영문, 최영두, "연관규칙을 이용한 데이터 분석에 관한 연구", 산업경영시스템학회지, 제 23권, 제 61집, pp. 115 - 126, 2000. 12.