

SOiVA기반 Learning UCC 메타데이터 참조 모델 설계

문남미*, 조태남**, 김경아***, 이은미***, 김형환****
 *호서대학교 벤처전문대학원 IT응용기술학과
 우석대학교, *명지대학, ****ETRI
 e-mail : mnm@hoseo.edu

Learning UCC metadata reference model design based on SOiVA

Nam-mee Moon*, Taenam Cho**, Kyong-Ah Kim***, Eun-mee Lee****, Kyunghwan Kim****
 *Dept of IT App. Tech, Hoseo University
 Woosuk University, *Myongji College, ****ETRI

요 약

Web2.0 시대를 맞아, 개인화(Personalize)와 사용자 중심(User Centric Method)의 참여형 서비스 시나리오에 기반하여 융합단말이 사용되어지고 있다. 대표적인 사용예가 UCC인데, 효과적인 지식공유와 적절한 정보를 추천하기 위해 효율적인 UCC 기반 콘텐츠의 사용의 참조(참조)가 가능하여야 한다. 본 논문에서는 UCC Learning의 효율적인 콘텐츠 양방향 서비스를 위하여, 참조 요소 관리를 할 수 있도록 데이터베이스를 설계 운영할 수 있는 방안에 관하여 연구하였다. 향후, 저작도구를 사용하여, UCC를 저작하는 한계를 극복하여, 다양한 콘텐츠방식 저작에 활용되어질 수 있도록 할 계획이다.

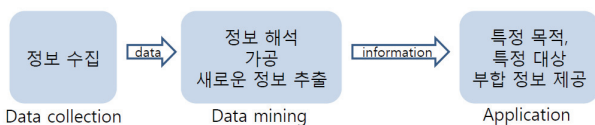
1. 서론

최근 디지털융합콘텐츠는 TV, 웹, 모바일, e-book 등 다양한 단말과 다양한 형태로 콘텐츠가 제작되고 있고 있으며, 사용자가 참여하여 다양한 콘텐츠 저작을 하여, SNS 통하여 배포하기도 하며 기존의 단방향 방식의 콘텐츠 소비와는 다른 양방향의 소비와 공급패턴을 가지고 있다.

학습의 범위도 다양해져서, 단순한 K12교육이 아니라, 평생교육차원의 다양한 교육이 사용자에게 의해서 제작되고 있다. 최근 선풍적인 사용을 보이고 있는 스마트 폰의 앱스토어는 소비자가 만든 어플리케이션이 공급되고 소비되는 좋은 예라 할 수 있다.

2. 추천시스템 개발 도구의 요구사항

추천 시스템은 온라인에서 이루어지는 개개인의 다양한 의사 결정을 지원하기 위해 다른 사람들의 직접적인 추천이나 사용 패턴에 대한 정보를 활용하는 온라인 의사 결정 지원 시스템들을 포괄적으로 의미한다.[1] 이를 다시 정의해본다면, 정보필터링을 사용하여 사용자에게 흥미로울 것 같은 정보아이템을 제공하는 시스템이라고 할 수 있겠다.



(그림 1) 추천시스템

2.1. 데이터수집

명시적인 수집방법과 잠재적인 수집방법으로 데이터수집을 나눠 표현 할 수있으나, UCC-Learning의 경우와 같이 일반적으로 SOiVA기반 콘텐츠는 두 가지 형태수집방법을 함께 병행 사용하게된다. 그러나,UCC의 경우 다양한 콘텐츠의 양은 기하급수적으로 늘어나고, 원하는 콘텐츠를 찾고자 할 경우 콘텐츠검색이나 자동추천과 같은 형태의 접근 방법이 있다. 콘텐츠 검색의 경우, 사용자는 찾고자하는 키워드를 찾아 검색엔진에 입력하고 해당검색결과를 기다려야한다.또한 검색결과로 제시된 프로그램 콘텐츠의 열거에서 다시 선택을 해야하는 절차를 갖게된다. 그러나 자동추천의 경우 이러한 단계를 줄일 수있다. 특히,UCC의 경우는 찾아진 콘텐츠가 또 다른 콘텐츠를 참조한 것이라면, 원래 참조한 것이 추천되는 것이 더 바람직 한 경우 또한 있다. 하나의 UCC콘텐츠는 저작행위 자체가 혼자 만든 것이 아니기 때문에, UCC형태 분석에 따라 참조유형또한 각각 다를 수 있다.

(표 1) UCC의 형태 분석 ([3]의 재수정)

분류	세분화	약어	예	유형	
형태	Generated	UGC	A	A가 고유한 창작에 의한 콘텐츠	순수창작물
	Modified	UMC	A+ a =A'	소스콘텐츠 A에 사용자의 아이디어 a를 덧붙인 A'는 A와 저작의도가 동일함	
	Recreated	URC	A+B=C	서로 다른 콘텐츠 AB를 조합하여 새로운 콘텐츠 C가 생성되었지만 저작의도는 AB와 다른 고유한 아이디어로 이루어짐	소스변형물

2.2. 메타데이터 설계

2.2.1 SOiVA 기반 디지털아이템 서비스 운영 모델 정의

SOiVA는 양방향 상호정보를 교환할 수 있게 하기 위해서는 기존의 디지털 아이템 자체에 하이퍼링크 기능, 소유권, SaaS 서비스 연계 기능, 가격 등의 표시를 가능하게 하여 상거래에서 사용자가 쉽게 디지털 아이템자체에서 정보를 보고 그와 관련된 서비스를 요청할 수 있으며 그 요청에 따라 내재된 하이퍼링크 기능 등의 다양한 기능을 이용해 다른 연계된 정보를 가질 수 있게 해야 한다[2] 이와 같은 운영모델에서 발생되어지는 트랜잭션은 크게 7가지 항목으로 분류할 수 있다.

- 디지털 생성배포(Distribution of Digital Creations)
- 생성된 디지털아이템에 고유번호할당(Allocation of Unique Number of Creation)
- 전자지불(Electronic Payments)
- 라이선싱 트랜잭션(Licensing Transaction)
- IPR정보 보고(Reporting IPR Information)
- 인증키(Certification Keys)
- 모니터링 사용(Monitoring Usage)

2.2.2 SOiVA기반 메타데이터 설계

SOiVA와 UCI의 메타데이터를 근간으로 하여 참조 요소를 추가하였다. 참조 요소에는 L_UCC가 참조한 콘텐츠소스에 관한 정보가 들어가게 된다.

(표 2) L_UCC의 메타데이터

	용어명	정의
1	Title	L_UCC의 제목
2	Idetifier	L_UCC의 유일한 전역 식별자
3	Description	L_UCC에 대한 부가 설명
4	Genre	L_UCC이 속하는 장르
5	Keyword	L_UCC를 나타내는 특정 단어
6	Language	L_UCC의 주 사용 언어
7	CreationDate	L_UCC의 생성일자
8	ProgramRef	L_UCC의 인스턴스를 가르키는 참조 식별자
9	FileFormat	L_UCC를 이루는 파일 포맷
10	Creator	L_UCC제작자
11	MediaDuration	L_UCC유효 기간, 유효 범위
12	Relation	L_UCC와 관련된 기타 정보 (URL, URI 등)
13	Rights	L_UCC저작권
14	Reference	L_UCC에 기여 및 공헌한 참조 아이템/아이템 제작자

참조 요소는 Children요소로 identifier, startTime, duration을 가지고 있다. identifier에는 참조 한 소스의 identifier를 명시해준다. startTime은 제작물에서 참조한 소스가 시작 되는 시간을 명시해준다. duration은 제작물에서 참조한 소스의 진행시간을 명시해준다.

(표 3) L_UCC의 참조 요소 구분

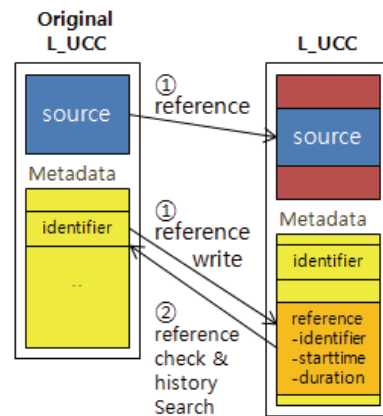
Diagram	
Children	<identifier><startTime><duration>n>
Source	<pre><xs:element name="참조"> <xs:complexType> <xs:sequence minOccurs="0"> <xs:element name="identifier"/> <xs:element name="startTime"/> <xs:element name="duration"/> </xs:sequence> </xs:complexType> </xs:element></pre>

참조요소의 활용방법은 다음과 같다.

- ① Original L_UCC로부터 특정 소스부분을 refernece 하는 경우, Original L_UCC의 identifier를 L_UCC의 참조 부분에 기록해주게 된다.
- ② L_UCC의 참조들은, 참조 check & History 검색을 하는데 활용 하게 된다.

```
if( L_UCC.참조.identifier == Original L_UCC.identifier )
  L_UCC ⇔ Original L_UCC Source
```

identifier의 참조를 이용해 참조한 소스를 알 수 있고, startTime와 duration을 이용해, 제작물이 어떠한 타임에, 어떠한 소스가 사용되었는지 확인 할 수가 있다.



(그림 2) 참조 요소를 이용한 메타데이터의 원리

여기에는 전제조건들이 따르게 된다. 공통된 부여규칙에 의해 L_UCC들의 identifier가 부여되어야 한다. 또 부여된 identifier들이 관리 될 수 있는 시스템 또한 필요하다.

2-3 참조 Storage

참조한 콘텐츠를 찾기 위해, 콘텐츠 메타데이터의 참조요소를 매번 확인 하는 것은, 다수의 콘텐츠 메타데이터에 접근하는 경우, 과부하를 가지고 올 수 있기 때문에 효율적인 방법이라 할 수 없다.

그래서 시스템상에 참조 Storage를 구축하고 참조 데이터만을 따로 관리하고, 활용 할 수 있도록 설계하였다.

참조 Storage에는 콘텐츠간의 연관관계가 저장되게 된다 (표 4)는 참조 DB의 구성이다. Sub_id는 DB에서 관리되는 주 키이다. Parent UCC ID에는 참조한 소스의 identifier가 들어간다. UCC ID에는 만들어진 UCC의 identifier가 들어간다. Modifier에는 참조의 형태를 넣어주게 된다.

Parent UCC ID와 UCC ID는 metadata Storage의 UCCidentifier 연관 관계를 이루어 줌으로써, 콘텐츠에 관한 정보를 가져 올수 있도록 한다.

(표 4) 참조 DB의 구성

Column Name	Data Type
Sub_ID	INTEGER
Parent UCC ID	VARCHAR(10)
UCC ID	VARCHAR(10)
Modifier	VARCHAR(5)

3. 협업필터링에 의한 추천시스템 설계

추천 시스템(Recommendation System)은 온라인 상에서 이루어지는 개개인의 다양한 의사결정을 지원하기 위해 다른 사람들의 직접적인 추천이나 사용패턴에 대한 정보를 활용하는 온라인 의사결정지원 시스템들을 포괄적으로 의미한다[4]. 내용기반 추천은 사용자가 평가했던 상품에 대한 특징정보와 다른 상품에 포함된 텍스트의 특징 정보의 유사도를 이용하여 필터링한다. 그러나 이 기법은 사용자가 이미 평가한 콘텐츠와 유사한 콘텐츠만을 제공하여 새로운 상품에 대한 사용자의 평가가 없는 경우에는 추천이 제한되는 문제점이 있다. 이와 같은 경우, UCC와 같이 방대한 분량의 콘텐츠는 추천되어지는 한계가 있다. 따라서 협업필터링을 사용하여 두단계를 거쳐 L-UCC를 추천할 수 있도록 한다.

단계1 : 피어슨 상관관계를 이용한 학습자 유사도 측정단계
 단계2: 참조모델을 활용하여, 추적되어진 참조 유형 평균으로부터의 편차에 따른 추천단계

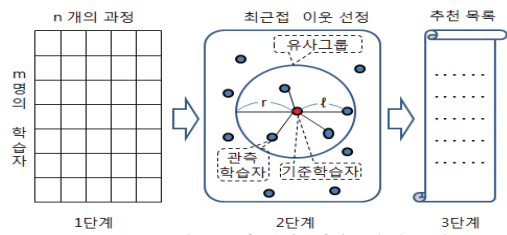
Pearson Correlation Coefficient는 학습자 A, B에 의해 공통적으로 평가된 아이템들의 평가치를 이용해서 (식 1)과 같이 계산한다. 여기서 A_i 와 B_i는 학습자 A와 B가 공통으로 평가한 아이템 i의 평가치를 뜻한다. 그리고 A, B 는 학습자 A와 B의 이용 가능한 평가치들의 평균값을 의미한다.

(표 5) 과목 X 학습자 매트릭스

고객X아이템 매트릭스

과목	과목 1	과목 2	과목 3	과목 n
학습자 1	R _{1,1}	∅	R _{1,3}	R _{1,n}
학습자 2	R _{2,1}	R _{2,2}	∅	R _{2,n}
.....	∅
학습자 m	R _{m,1}	∅	R _{m,3}	R _{m,n}

$$w(A, B) = \frac{\sum_{i=1}^n (R_{A,i} - \bar{R}_A)(R_{B,i} - \bar{R}_B)}{\sqrt{\sum_{i=1}^n (R_{A,i} - \bar{R}_A)^2 \sum_{i=1}^n (R_{B,i} - \bar{R}_B)^2}} \quad (식1)$$

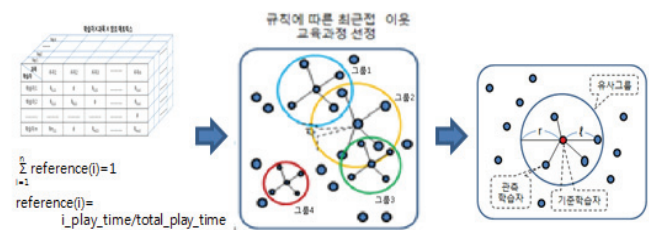


(그림 3) 최근접 이웃 선정단계

그러나, UCC의 경우에는 참조되어진 것이 하나의 동영상이라해도, 실질적으로는 하나 이상의 것일 수있다. 즉, 하나의 교육을 위한 SCO(Sharable Content Object)를 위해서 다양한 SCO와 결합할 수 있는 것이다. 새로운 SCO를 만들기 위해서, 결합되어진 SCO와 Asset에 가중치를 두어 표준편차를 이용해서 최근접 이웃을 선택하게 함으로써 개선된 CF를 제안한다.

학습자X과목 X 참조 매트릭스

과목	과목1	과목2	과목3	과목n
Rep 1	R _{1,1,1}	∅	R _{1,3,1}	R _{1,n,1}
Rep 2	R _{2,1,1}	R _{2,2,1}	∅	R _{2,n,1}
.....	∅
Rep k	R _{k,1,1}	∅	R _{k,3,1}	R _{k,n,1}



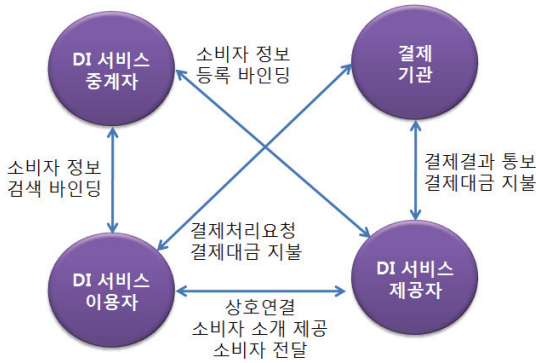
(그림 4) 참조되어진 정보분석에 따른 그룹선정과 최근접이웃선정 : 개선된 CF추천과정

즉, 참조되어진 정보를 고려하여 L-UCC의 그룹을 선정하고 다시 그 그룹내에서 사용자의 최근접 이웃을 찾아 과정을 추천할 수 있도록 한다. 이와 같은 L-UCC콘텐츠의 저작권을 고려하여, 판매하게 된다면 이는 U-Commerce 비즈니스모델에 근거하여 판매되어진다. U-Commerce의

Value(U-Value)는 다음과 같이 표현할 수 있다.[2] L-UCC도 이와 같은 원칙에 따라, 학습자에게 판매가능하다.

$$U-Value=f(\text{seamlessness, convenience, privacy})$$

특히, P2P를 제공하는 U-Commerce환경은 Seamlessness성격으로 정보를 수집하기 좋은 환경이기 때문에 심각한 프라이버시 문제가 초래될 수 있다. 이 경우, 정보 생성자인 end user가 최대 정보 소유자이며, 사업자는 필요 정보만을 소유하는 형태로 End user 중심으로 아키텍처를 설계하여 프라이버시를 보호 할 수 있다. 이와 같은 u-commerce환경은 거래 행위에 있어 각 개인에게 상시적이고 일상적인 정보처리행위를 요구할 수 있다. 즉, 사용자 개인의 디바이스는 다양한 거래를 지원하는 강력한 도구로서 진화하게 될 것이다. 향후, 이와 같은 진화는 Seamlessness가 지켜지면서도 사용자의 프라이버시가 지켜줄 수 있는 지불 시스템이 필요로 되어 진다. (그림5)는 디지털아이템 서비스이용자와 제공자, 그리고 중계자간에 서비스 시나리오를 보여준다.



(그림 5) 서비스 시나리오 액터 구성

L-UCC의 경우, DI서비스제공자는 학습콘텐츠의 Asset이나 SCO를 제공한 다수가 될 수 있으므로, 이는 데이터베이스에 기록되어있는 참조모델을 참조하여, 거래 행위나 정보처리를 요구할 수 있다.

4. 결론 및 향후 연구

지금까지 참조모델을 사용하여, History 검색을 위한 L-UCC의 메타데이터 및 DB 설계와 History 시스템 구조에 대해서 정의, 설계 하였다.

History 검색을 통해서 다양한 콘텐츠 제공 출처가 명확해 질 수 있어, 신뢰도와 상거래에 저작권에 관련된 문제해소에 도움이 될 수 있으리라 기대한다.

본 연구는 저작도구를 사용하여, 저작되어진다는 한계를 가지고 설계를 하였으나, 향후 폭넓은 구현 설계에 근거하여 보다 효율성있는 시스템 설계를 할 필요가 있다고 판단한다.

참고문헌

[1] ATSC, ATSC standard : ATSC interaction channel protocols, www.atsc.org, 2006
 [2] 문남미의, "SOiVA를 활용한 U-Learning 비즈니스모

델설계”, 한국정보처리학회 춘계학술발표대회논문집 제16권
 [3] “You! UCC”, 마케팅사관학교&김영한, 랜덤하우스, 2007.2 p.30
 [4] Segaran,Toby,Programming Collective Intelligence, Oreilly & Associates Inc,2007