

웹 서비스를 지원하는 TV-Anytime 개인화 서비스 시스템의 설계 및 구현

이종설*, 이석필*, 신사임*
*전자부품연구원
leejs@keti.re.kr

Design and Implementation of a TV-Anytime System for Personalized service supporting the Web Service

Jong-Seol Lee*, Seok-Pil Lee*, Sa-Im Sin*
*KETI

요 약

본 논문에서는 TV-Anytime Forum에서 정의된 양방향 환경에서의 사용자 정보에 기반한 맞춤형 콘텐츠 및 광고 서비스를 설계 및 구현한다. 구현된 시스템은 TV-Anytime Forum 표준을 만족하며, 콘텐츠의 로케이션 리졸빙, 멀티미디어 콘텐츠의 추천 등을 지원한다. 이를 위하여 메타데이터 저장 모듈, 개인화 TV 에이전트, 사용자 콘텐츠 사용정보 생성기 등을 구현했으며, 클라이언트 단말에서는 SOAP 오퍼레이션을 통해 콘텐츠 정보를 수집한다. 클라이언트에서는 사용자 정보를 생성 및 관리하며 이를 바탕으로 사용자에게 적합한 콘텐츠 및 광고 정보를 추천한다.

1. 서론

디지털 방송환경의 발전에 따라 사용자는 방송 콘텐츠 뿐 만이 아닌 콘텐츠의 부가정보에 대한 서비스가 가능해졌으며 이를 통한 리치 메타데이터 서비스가 가능해졌다. 디지털방송에서는 PSI 정보 등을 바탕으로 콘텐츠의 타이틀, 방송 시간, 줄거리 등 간단한 부가 정보의 서비스가 가능해졌다. 이와 같은 정보는 매우 제한된 정보로 보다 풍부한 정보의 제공 및 부가 정보를 통한 서비스들이 요구되어 졌으며 대표적인 부가 정보를 위한 표준화 단체가 TV-Anytime Forum이다. TV-Anytime Forum은 디지털 저장 매체를 이용한 방송 콘텐츠 및 양방향 네트워크 콘텐츠를 위한 검색, 저장, 메타데이터 생성 방법, 보안 등에 대한 표준화를 목적으로 설립 되었다. 이를 통하여 사용자는 기본적으로 콘텐츠에 대한 부가 정보 및 사용자 정보를 활용한 서비스가 가능해졌다. 또한, 광고에 대한 부가 서비스가 가능하다. TV-Anytime Forum은 2001년 8월에 ver.1.1 Metadata 규격을 완성하였으며 2005년 6월 Phase 2에 대한 표준을 완료하고 ETSI(유럽전기통신표준협회)의 방송 및 온라인 서비스 표준으로 채택 되었다. 현재는 관련 기술을 바탕으로 IPTV 및 DMB 분야 등에서 확장 중이다. 본 논문에서는 TV-Anytime Forum 표준에 기반 하여 사용자 정보에 기반 한 맞춤형 콘텐츠 및 광고 서비스를 위한 시스템을 설계 및 개발한다.

2. TV-Anytime 표준

(1) 메타데이터

TV-Anytime 메타데이터의 종류는 크게 4부분으로 나눌 수 있다. 첫째로 Content description 메타데이터로 콘텐츠에 대한 기술정보를 나타내는 메타데이터다. 여기에는 Description, Audio/Video Information, Program Information, Group Information, Media Review DS에 대해서 정의한다. 콘텐츠의 기술정보로는 프로그램 제목, 장르, 줄거리 개요, 등장인물 정보 등이 해당된다. 둘째는 프로그램에 관련된 프로그램 방송시간, 채널정보, 다운로드 위치, 프로그램 개체의 이용 규정 등 로케이션 정보와 서비스 정보를 나타내는 Instance Description 메타데이터다. 셋째는 한 콘텐츠 안에서 하이라이트나 키 프레임을 나타낼 수 있는 Segmentation 메타데이터며, 마지막으로 사용자의 취향이나 습성을 나타내는 Consumer 메타데이터로 사용자의 콘텐츠 사용 내역에 대한 정보인 Usage History와 사용자가 직접 기입한 사용자의 선호 정보인 User Preference로 나누어진다.

(2) Location Resolution

TV-Anytime 시스템에서 사용자는 방송망 또는 리턴채널을 통해 전송된 ECG 정보 또는 트레이일러를 통해 멀티미디어 콘텐츠에 대한 정보를 획득하거나, 콘텐츠 검색 인터페이스를 통해 콘텐츠를 검색하여 콘텐츠 선택할 수 있다. TV-Anytime에서는 이와 같은 콘텐츠의 획득 절차를 위해 Content Reference Identifier (CRID)를 정의하였다. CRID는 사용자가 선택한 콘텐츠가 정확히 어떤 것인지를 가리키는 일종의 식별자이다. 선택된 CRID는 위치 분석(Location Resolution)을 통해 콘텐츠의 저장된 위치, 또는

방송 시간 등을 나타내는 로케이터를 얻을 수 있고, 이 로케이터를 이용해 사용자는 선택한 콘텐츠를 획득 할 수 있다. 하나의 CRID는 위치 분석과정에서 여러 가지의 CRID들로 나뉘어 질 수 있는데, 예를 들어 사용자가 선택한 콘텐츠인 사극 '제국의 아침'이 속편들을 포함한 타이틀인 경우 위치 분석 과정에서 최초 CRID는 각 속편들 하나 하나의 CRID들로 나뉘어 진다. 이와 같은 CRID를 표현하는 형식은 다음과 같다.

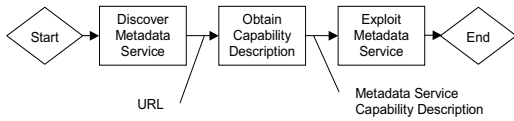
CRID://<authority>/<data>

authority는 콘텐츠에 대해 저작권을 가진 콘텐츠 제공자의 도메인 이름으로 이루어지며 동일 도메인 이름을 가진 여러 개의 권한이 있을 경우 추가 될 수 있다. authority는 아래와 같은 형식으로 표현 될 수 있다.

<DNS name><name_extension>
따라서 CRID는 아래 형식으로 표현된다.

CRID://<DNSname>
<name_extension>/<data>

(3) 양방향 환경에서의 메타데이터 검색 및 전송



(그림 1) 웹 서비스를 통한 TV-Anytime 서비스의 이용 절차

W3C에서 규격화 하고 있는 SOAP은 현재 웹 서비스에서 가장 핵심인 프로토콜이다. 차세대 웹 서비스는 기본적으로 UDDI 표준을 이용하여 서비스를 검색하고, WSDL 표준을 이용하여 해당 서비스의 내용을 파악한 뒤 SOAP 표준을 통해 실제적인 웹 서비스 메시징 작업을 수행하게 된다. 이를 위해 TV-Anytime Forum에서도 양방향 메타데이터 전송을 위해 WSDL 및 SOAP 대한 스키마를 정의하였으며 이를 표준화하고 있다.[1]

● get_Data

get_Data 메소드는 사용자가 서버상의 TV-Anytime 메타데이터에 대해서 검색 질의를 위한 인터페이스를 제공한다. get_Data 메소드를 이용하여 사용자는 타이틀, 장르, 배우, 등과 같은 검색 값을 통해 원하는 메타데이터 테이블을 추출한다.

● submit_Data

submit_Data 메소드는 사용자 메타데이터인 콘텐츠 사용 정보(Usage history)를 익명(Anonymous) 상태로 서비스 제공자 서버에 전송한다.

● upload_Personal_Data

submit_Data와는 달리 사용자의 모든 정보를 비익명(non-Anonymous) 상태로 전송한다.

● clear_Personal_Data

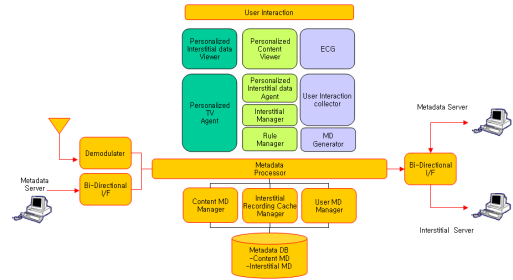
서비스 프로바이더에서 관리하는 사용자 정보를 삭제 및 관리 하기 위한 오퍼레이션이다.

(4) Interstitial Metadata

광고 메타데이터의 저장, 재생, 관리를 나타내는 Interstitial Metadata는 크게 재생 규칙을 표현하는 Rules, 광고의 저장을 나타내는 RecordingCache, 광고의 대체 재생을 나타내는 InterstitialTable로 구성된다. 이를 통해 단말은 광고의 저장 및 교체 재생이 가능하다.

3. 개발 내용

TV-Anytime 메타데이터는 방송 스트림에 포함되거나 또는 SOAP 프로토콜을 통해 전송 받을 수 있다. 이와 같이 전송된 메타데이터를 처리하기 위하여 (그림 2)와 같이 시스템을 구성하였다.

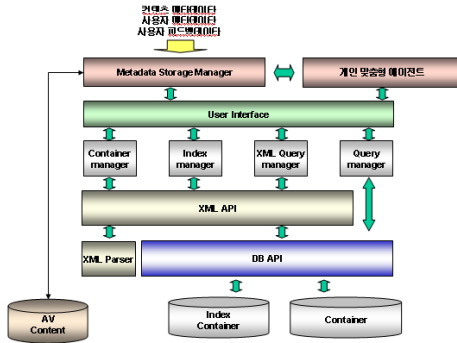


(그림 2) 시스템 구조

(1) 메타데이터 저장 관리 모듈

TV-Anytime 메타데이터는 XML 문서의 형태로 전송되어진다. XML 문서는 비분할 형태(XML 문서 전체가 기본 단위)와 분할 형태(XML 문서의 엘리먼트가 기본 단위)로 컨테이너에 저장되어 진다. 분할 형태로 저장되는 엘리먼트는 Index manager를 통해 설정/관리되어진다. 기본적으로 자주 참조되어지는 엘리먼트를 설정한다. 문서의 전체 내용은 비분할 형태로 컨테이너에 저장되어진다. 검색에 있어서는 분할 형태가 더 빠르지만 문서의 로딩에 있어서는 비분할 형태가 더 나은 성능을 보인다. 따라서 참조되어지는 빈도를 고려하여 설정한다. Database API는 가장 하부의 물리적 저장소인 Index Container와 Container를 다루기 위한 기본적인 데이터베이스 연산들의 집합이다. Index Container에는 Index 저장 관리하며, Container에서는 메타데이터 내용을 저장한다. 하나의 Container는 하나의 Index Container를 갖는다. XML API는 XML파서를 포함하며 XML과 관련된 기본적인 작업을 수행한다. XQuery manager에서는 XQuery 질의 언어를

분석하고 처리하기 위한 모듈이다. XQuery는 전체 문서, 문서의 일부분, 문서에 포함된 하나 이상의 노드들이 가지는 값에 대한 검색이 가능하다. 또한 문서 질의에 의해 반환되는 값들을 조작하거나 변환하는 것이 가능하다. XQuery는 XPath를 사용하여 문서나 각 노드들을 표현한다. Query Manager는 non-XQuery 질의를 위한 모듈이다. XQuery에 비하여 XML 처리 단계를 필요 하지 않으므로, 그 처리 속도가 빠르다.



(그림 3) 메타데이터 저장 관리 모듈

(2) 메타데이터 인덱스 관리

과실 모듈을 통해 구조 분석된 메타데이터는 메타데이터 객체 관리자를 통해 저장된 이후 인덱스 관리자에게 전달된다. 인덱스 관리자에게는 메타데이터의 검색속도와 관리의 편리성을 위해 그림과 같이 역 파일 인덱스 형태로 저장된다. 첫째로 (그림 4)와 같이 기본적으로 필요한 값들에 대한 정보를 저장한다. 이때 CONTENTID가 각 콘텐츠를 대표한다. 콘텐츠에 대한 검색을 위해 Title, Genre, Cast, Keyword등에 대한 DS 값과 CONTENTID에 대해서 역파일 인덱스를 생성 관리한다. 마지막으로 개인화 정보를 위해 UserID와 개인 선호 정보 항목에 대해서도 역파일 인덱스를 생성 관리한다. TV-Anytime 메타데이터는 Table 단위로 구분되어며, Container manager와 Index manager를 통해 각 Table은 하나의 Container에 저장되며 각 Container에 대한 Index는 Index Container에 저장되어진다.

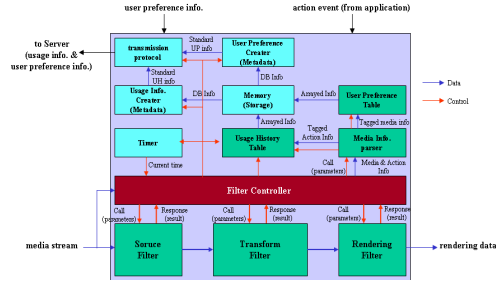
ContentID	CRID	Genre	Title	Keyword	Channel	Time
CRID	ContentID's	Keyword	ContentID's			
Genre	ContentID's	Channel	ContentID's			
Title	ContentID's	Time	ContentID's			

(그림 4) 메타데이터 인덱스 관리자의 역파일 인덱스 구조

(3) 사용자 콘텐츠 사용정보 생성 모듈

사용자 콘텐츠 사용정보 생성 모듈은 사용자가 콘텐츠 또

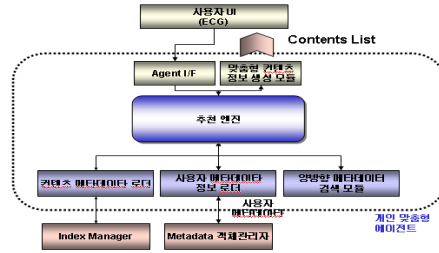
는 디바이스를 사용한 정보이다. 그림과 같이 사용자 콘텐츠 사용정보 생성 모듈은 콘텐츠에 대한 Filter controller와 연결되어 Timer를 통한 시간정보를 바탕으로 콘텐츠 사용 정보를 생성하게 된다.



(그림 5) Usage history 생성 모듈

(4) 개인화 TV 에이전트(Personalized TV Agent)

개인화 TV 에이전트는 그림과 같이 콘텐츠 메타데이터 로더, 사용자 메타데이터 정보 로더, 양방향 메타데이터 검색 모듈, 추천엔진, Agent Interface, 맞춤형 콘텐츠 정보 생성 모듈 등으로 구성 되어 있다.



(그림 6) Personalized TV Agent

콘텐츠 메타데이터 로더는 메타데이터 저장/관리 모듈의 인덱스 매니저를 통해 콘텐츠 메타데이터의 특정 벡터 값을 읽어 들이며, 사용자 메타데이터 정보 로더는 메타데이터 저장/관리 모듈의 메타데이터 객체 관리자를 통해 추천 작업을 수행 할 사용자 정보를 읽어 들인다. 양방향 메타데이터 검색 모듈은 양방향 환경을 통해 메타데이터 서버를 통한 메타데이터 추천 작업을 수행한다. 추천엔진은 콘텐츠 메타데이터 로더 및 사용자 메타데이터 정보 로더를 통해 읽어 들인 콘텐츠 메타데이터 벡터 값들과 사용자 정보를 바탕으로 추천 작업을 수행하게 된다. 알고리즘은 유사 선호도를 이용한 계산방법을 수행한다. Agent Interface를 통해 추천 에이전트는 각 사용자의 선택에 따라 추천 시점, 양방향 추천 서비스 검색 설정 등이 가능하다. 맞춤형 콘텐츠 정보 생성 모듈은 메타데이터 추천 에이전트를 통해 생성된 콘텐츠 리스트를 사용자 UI를 통해 사용자에게 제공되어진다. 이때 사용자에게 제공되어질 콘

텐츠들의 선별 작업을 수행하게 된다.

추천엔진은 Container에 저장된 콘텐츠 정보와 사용자 선호 정보를 이용하여, 일정한 추천 알고리즘에 의해 사용자가 선호할 만한 프로그램들을 자동으로 선정하여 사용자에게 추천하게 된다. 다음은 추천엔진에서 사용자의 선호 정보와 콘텐츠 정보를 이용하여 사용자가 선호할만한 프로그램들을 자동으로 추천하는 과정을 단계별로 설명한 것이다.

-Step 1 : 항목별로 사용자 선호도 추출

$$p = (w_1, w_2, w_3, \dots, w_n)$$

-Step 2 : 콘텐츠 정보로부터 항목별(Title, Channel, Genre, Cast, Keyword, Time등)로 프로그램들의 항목값을 추출

$$C = (c_1, c_2, c_3, \dots, c_n)$$

- Step 3: 유사도 값을 계산한다.

$$Similarity(P, C) = \frac{\sum_{i=1}^n w_i c_i}{\sum_{i=1}^n w_i}$$

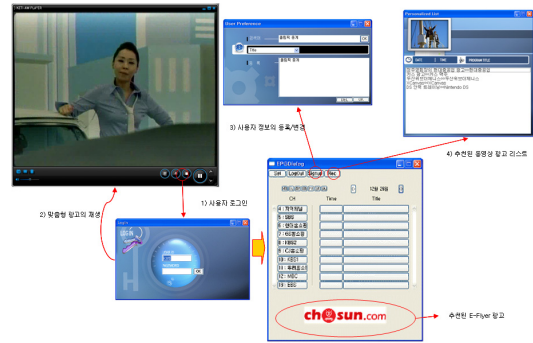
- Step 4: 추천 후보 프로그램들을 각각의 유사도 값에 따라 정렬

- Step 5 : 상위 N개의 프로그램들만 사용자에게 추천

4. 결론

본 논문에서는 사용자의 정보에 기반 한 콘텐츠 및 광고의 맞춤형 서비스를 위한 에이전트를 설계 및 구현 하였다. 본 논문에서 개발된 맞춤형 서비스 기술은 디지털 방송, 케이블방송, 위성방송, IPTV, DMB 등 어떠한 전송매체를 통하더라도 서비스가 가능하다. 개인 맞춤형 에이전트 기술은 무수히 많은 채널 및 콘텐츠의 홍수 속에서 사용자에게 보다 유익하고 편리한 인터페이스를 제공할 것이며, 서비스 제공자에게는 새로운 비즈니스 영역을 제공 가능하다.

본 연구에서 구현한 시스템은 단방향과 양방향 서비스를 모두 제공할 수 있는 시스템으로 개인의 취향에 따른 맞춤형 방송 서비스 분야에 주는 영향이 클 것으로 예상된다.



(그림 6) 맞춤형 콘텐츠 및 광고 뷰어

참고문헌

- [1] Requirements and Scenarios for the Bi-directional Transport of Metadata, TV150r1The TV-Anytime Forum<http://www.tv-anytime.org>
- [2] Simple Object Access Protocol (SOAP) 1.1, W3C Note, 8 May 2002D. Box, et. al. <http://www.w3.org/TR/2000/NOTE-SOAP-20000508/>
- [3]TV-Anytime Requirements Series: R-1, TV035r6The TV-Anytime Forum<http://www.tv-anytime.org>
- [4] TV-Anytime Specification on Content Referencing, SP004v1.2The TV-Anytime Forum<http://www.tv-anytime.org>
- [5] TV-Anytime Specification on Metadata, SP003v1.3The TV-Anytime Forum<http://www.tv-anytime.org>
- [6] TV-Anytime Specification on Metadata Protection, SP007v1.0The TV-Anytime Forum<http://www.tv-anytime.org>
- [7] TV-Anytime Specification on System, SP002v1.3The TV-Anytime Forum<http://www.tv-anytime.org>
- [8] Web Services Description Language, Version 1.1, W3C Note 15 March 2001E. Christensen, F. Curbera, G. Meredith, S. Weerawarana<http://www.w3.org/TR/2001/NOTE-wsdl-20010315>