

TV-Anytime 메타데이터 관리 시스템

박종현⁰ 김병규 강지훈
충남대학교 컴퓨터학과
{ jhpark, yourovin, jhkang }@cs.cnu.ac.kr

TV-Anytime Metadata Management System

Jong-Hyun Park⁰, Byung-Kyu Kim, & Ji-Hoon Kang
Dept. of Computer Science, Chungnam National University

요 약

TV-Anytime 은 국제 표준으로 자리잡아가고 있는 차세대 디지털 방송용 메타데이터로서 사용자가 직접 원하는 방송을 검색하고 선택하여, 선택한 멀티미디어 데이터를 저장하였다가 시간과 장소에 구애받지 않고 사용할 수 있도록 하는 것을 목표로 하고 있다. TV-Anytime 은 XML 스키마로 정의되어 있으며, 방송이라는 특성상 그 크기에 있어서 대용량이다. 본 연구에서는 이러한 TV-Anytime 의 특성에 맞는 메타데이터 관리 시스템을 제안한다. 대용량의 문제를 해결하기 위하여 데이터베이스를 사용하며, TV-Anytime 스키마를 분석하여 그에 적합한 스키마를 설계한다. XML 형식의 TV-Anytime 메타데이터에 대한 질의를 위하여 XQuery 를 사용한다. 이는 데이터베이스의 저장 형식에 무관하게 질의하는 것을 가능하게 만들어 주며, 디지털방송 환경에서 시스템 간의 상호 운용성이 확보될 수 있다. 또한, XQuery 는 정보검색 기능을 가지고 있어서 향후 콘텐츠 기반 정보검색을 포함하는 등 보다 풍부한 의미를 갖는 질의를 사용할 수 있을 것이다.

1. 서론

TV-Anytime 기술은 차세대 디지털 방송용 메타데이터 국제 표준으로 자리잡아 가고 있는 기술로 점점 대용량화 되고 저렴해지는 저장장치를 이용하여 사용자가 원하는 멀티미디어 데이터를 검색하여 저장하였다가 시간과 장소에 구애 받지 않고 사용할 수 있도록 하는 기술이다.[6,7] 이러한 TV-Anytime 메타데이터는 단일의 스키마[14]를 따라 XML[10]을 사용하여 기술하고, 방송이라는 특성상 메타데이터의 양이 대용량이다.

본 논문에서는 이러한 TV-Anytime의 특성에 맞는 메타데이터 관리 시스템을 제안한다. 먼저, 대용량의 문제를 해결하기 위하여 데이터베이스를 사용한다. TV-Anytime은 단일 XML 스키마로 정의되어 있다. TV-Anytime 메타데이터를 데이터베이스에 저장하기 위해서는 XML 데이터를 데이터베이스로 변환하여야 하므로 TV-Anytime에 적합한 데이터베이스 스키마를 정의한다. 이때 단일 XML 스키마를 사용한다는 점을

고려하여, 다수의 XML 스키마에 기반을 둔 XML 데이터를 데이터베이스에 저장하기 위하여 데이터베이스 스키마를 자동으로 자동으로 생성하는 방법 대신, TV-Anytime 스키마를 분석하여 그에 적합한 스키마를 직접 설계한다.

XML 데이터인 TV-Anytime 메타데이터에 대한 질의를 위하여서는 XML 표준 질의인 XQuery[9]를 사용한다. 이는 데이터베이스의 저장 형식에 무관하게 질의하는 것을 가능하게 만들어 준다. 디지털방송 환경에서 고객의 PDR(Personal Digital Reorder)에서도 XQuery를 지원함으로써 시스템 간의 상호 운용성이 확보될 수 있다. 또한, 아직 확정되지는 않았지만 XQuery는 정보검색 기능을 가지고 있어서 향후, 단순 텍스트 뿐만 아니라 콘텐츠 기반 정보검색을 포함하는 등 보다 풍부한 의미를 갖는 질의를 사용할 수 있을 것이다.

본 논문의 구성은 다음과 같다. 2장에서는 TV-Anytime기반 방송 시스템과 TV-Anytime 메타데이터의

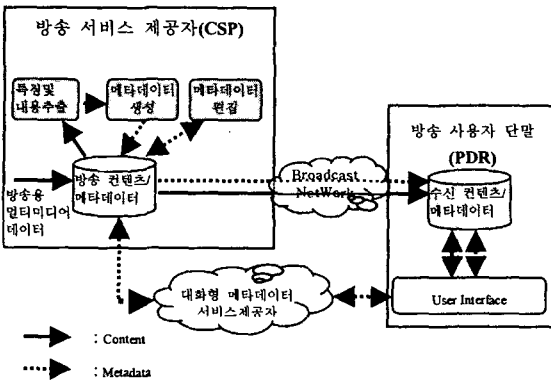
* 이 연구는 BK21 충남대학교 정보통신인력양성사업단의 지원을 받았다.

구조, 그리고 메타데이터 검색을 위한 질의어인 XQuery에 대하여 설명하고 있다. 3장에서는 TV-Anytime 메타데이터 관리 시스템과 시스템의 구성 요소인 저장 엔진과 XQuery엔진을 설명하고 있으며, 4장에서 향후 연구 방향과 결론을 맺는다.

2. TV-Anytime기반 방송 시스템

2-1. TV-Anytime기반 방송 시스템의 구조

TV-Anytime 메타데이터 기반의 방송 시스템은 [그림 1]과 같이 구성된다.[15]



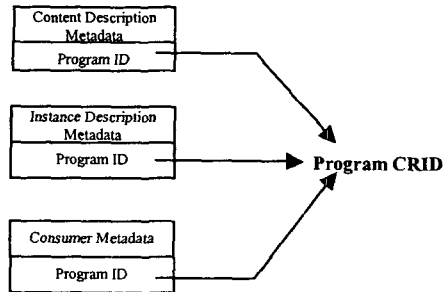
[그림 1] 메타데이터 기반 방송 서비스 시스템

메타데이터 기반 방송서비스 시스템은 크게 방송 서비스 제공자(Content Service Provider)와 서비스를 사용하는 방송 사용자 단말(Personal Digital Recorder)로 구성되고, 방송 사용자 단말은 다시 모바일 폰과 같은 하위의 방송 사용자 단말(Sub-PDR)로의 데이터 전송이 가능하다. 방송서비스 제공자는 방송용 멀티미디어 데이터에서 특징 및 내용을 추출하여 메타데이터를 생성하고, 추출된 메타데이터에 편집이 필요한 부분은 메타데이터 편집기를 이용하여 방송용 멀티미디어 데이터에 대한 메타데이터를 완성한다. 이때 방송용 메타데이터는 TV-Anytime 스키마에 정의되어 있는 정보를 추출하고 생성한다. 이렇게 모든 방송용 멀티미디어 데이터마다 각각 생성된 메타데이터는 저장되고, 필요에 따라 방송망을 통하여 방송 사용자 단말로 전송된다. 이때 기존 공중파 방송 서비스 시스템과는 달리 방송망(유.무선에 관계없이)으로 전송되는 데이터는 방송용 멀티미디어 데이터뿐 아니라 메타데이터 역시 함께 전송된다. 방송 사용자 단말은 수신된 방송용 멀티미디어 데이터와 메타데이터를 저장장치에 저장하여 두었다가 언제 어디서든지 사용자의 요청이 있을 경우 제공한다. 사용자가 원하는 멀티미디어 데이터를 찾고자 할 경우 사용자는 메타데이터 검색을 통하여 멀티미디어 데이터를 검색하고 검색한 메타데이터에 해당하는 멀티미디어 데이터를 방송서비스 제공자에게 요청할 수 있다.

이러한 디지털 방송서비스 시스템에서 콘텐츠 검색, 사용자 기록/선호도 관리, On-Demand 서비스 등 새로운 부가 서비스를 원활하게 지원하기 위해서는 방송용 메타데이터는 반드시 필요하다. 그러므로 메타데이터를 효율적으로 관리하는 것은 매우 중요하다.

2-2. TV-Anytime 메타데이터의 구조

TV-Anytime 메타데이터 규격에서는 XML Schema언어를 사용하여 그 구조를 정의 하는 한편 각 엘리먼트의 속성에 대한 의미(Semantics)를 함께 규정하고 있다. TV-Anytime 규격에 따르면, 메타데이터는 크게 멀티미디어 데이터 자체를 기술하는 메타데이터, 프로그램 개체를 기술하는 메타데이터 및 사용자 측면의 정보를 기술하는 메타데이터로 나눌 수 있다.



[그림 2] Program CRID를 참조하는 메타데이터의 구조

- 콘텐츠 기술 메타데이터(Content description metadata)
 - 프로그램 제목, 장르, 줄거리 개요, 등장인물 정보 등
- 개체 기술 메타데이터(Instance description metadata)
 - 프로그램 위치(방송시간, 채널정보), 프로그램 개체의 이용규정 등
- 사용자 메타데이터(Consumer metadata)
 - 이용내역(usage history), 사용자 선호도 정보 등

이들 메타데이터는 [그림 2]와 같이 콘텐츠를 참조하기 위한 식별자인 CRID(Content Referencing Identifier)를 매개로 동일한 콘텐츠에 대한 정보를 기술하게 된다.[16]

2-3. XQuery

현재 인터넷 표준화 단체인 IETF와 W3C는 공동으로 XML 데이터의 검색을 위한 질의어로 XQuery를 국제 표준으로 제정 중에 있다. XQuery는 기존의 XML을 위한 질의어인 XQL[11]이나 XML-QL[12]등과 같은 XML을 위한 질의어의 장점을 모두 포함하고 있는 질의어이다. 본 논문에서는, TV-Anytime 메타데이터의 저장 형식에 무관하게 질의할 수 있고 상호 운

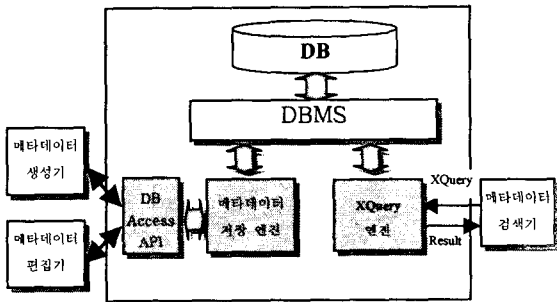
용성을 유지하기 위하여 TV-Anytime 메타데이터의 검색을 위한 질의어로 XQuery를 사용한다.

3. TV-Anytime 메타데이터 관리 시스템의 설계

3-1. 시스템 구성

TV-Anytime 메타데이터 관리 시스템은 TV-Anytime 스키마의 구조에 적절하게 정의된 데이터베이스 스키마에 메타데이터를 효율적으로 저장하고 관리하는 역할을 담당한다. TV-Anytime 메타데이터는 방송용 메타데이터의 양이 대용량이라는 문제점을 효과적으로 처리하기 위하여 텍스트 파일 기반의 메타데이터 저장 관리 방식보다는 데이터베이스 기반의 저장관리 시스템을 사용한다. 또한 시스템의 상용화에 대비하여 보다 고급의 검색서비스를 할 수 있도록 하면서도 방송 네트워크 상의 시스템 간 상호 운용성을 확보할 수 있는 방안으로 XQuery 엔진을 두어 편집기에서 검색을 위하여 또는 다른 응용에서 요구하는 XQuery를 처리할 수 있도록 한다.

[그림 3]은 메타데이터 기반 방송서비스 시스템에서 메타데이터 관리 시스템의 기능과 역할을 보여주고 있다. 메타데이터 관리 시스템은 크게 메타데이터의 저장과 관리를 담당하는 저장엔진과 XQuery 질의어를 입력 받아 처리하는 XQuery엔진으로 구성된다.

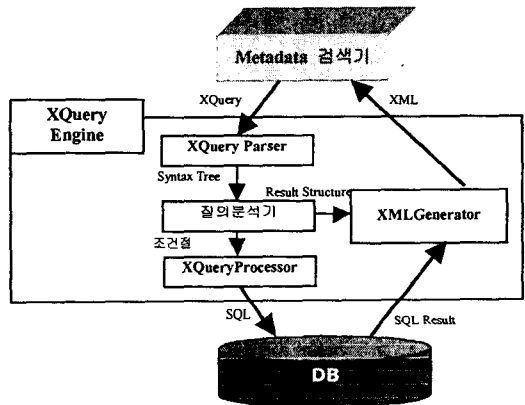


[그림 3] TV-Anytime 메타데이터 관리 시스템

메타데이터가 데이터베이스에 저장됨에 따라 메타데이터 생성기와 편집기는 데이터의 입력과 출력을 위한 데이터베이스 접근 방식이 필요하며, 이를 위하여 TV-Anytime 메타데이터 관리 시스템에서는 데이터베이스 접근 API를 제공하고 있다. 메타데이터 저장엔진은 편집기가 데이터베이스 접근 API를 통하여 요청하는 TV-Anytime 메타데이터를 데이터베이스에 효율적으로 저장하는 기능을 담당한다. 서버 사용자는 필요한 메타데이터를 검색하기 위하여 메타데이터 검색기를 사용한다. 사용자의 질의는 XQuery 질의어로 변환이 되며, XQuery 질의는 XQuery 엔진으로 보내어지며, 엔진은 메타데이터 데이터베이스를 이용하여 질의를 처리하여 그 결과를 XML 형태로 변환하여 메타데이터 검색기로 돌려준다.

3-2. 메타데이터 검색용 XQuery엔진

XQuery 엔진은 방송 서비스 제공자(Content Service Provider) 시스템의 하나의 모듈로서 존재하며, 방송 서비스 제공자는 방송용 멀티미디어 메타데이터의 검색을 위하여 메타데이터 검색기를 사용한다. 이때 질의는 내용 기반 질의와 같은 고급의 검색 기능을 지원하는 것이 궁극적인 목표일 것이다. 주어진 질의는 XQuery 질의로 변환이 되며, XQuery 질의는 XQuery 엔진으로 보내어지며, 엔진은 질의를 처리하여 그 결과를 XML 문서의 형태로 메타데이터 검색기로 돌려준다



[그림 4] XQuery 엔진의 구조

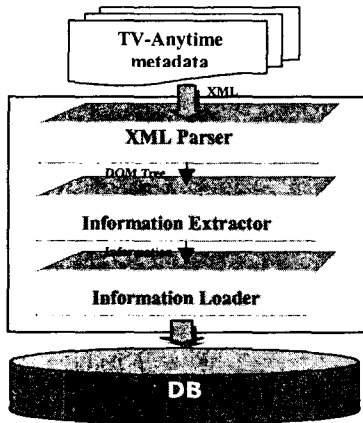
XQuery 엔진의 구조는 다음 [그림 4]와 같은 구조를 가진다. 메타데이터 검색기는 사용자가 질의 하는 질의어를 XQuery로 변환하여 XQuery엔진에 질의 한다. XQueryParser는 XQuery 질의어를 구문 분석하여 XQuery 구문 분석 구조(Syntax Tree)를 생성한 후 질의 분석기에게 제공한다. 이때 결과값의 구조 정보를 위해 return절에 있는 반환되어야 할 XML문서의 구조 정보는 XMLGenerator에게 제공하고 검색을 위해 필요한 조건절 부분은 XQueryProcessor에게 제공한다. XQueryProcessor는 XQuery의 Syntax Tree를 입력으로 XQuery에 기술되어 있는 표현(Expression)과 기능(Function)들을 얻어내어 데이터베이스에 질의 가능한 SQL로 변환한다. 변환된 SQL은 데이터베이스에 질의 하여 적절한 결과를 얻는다. 이렇게 얻어진 SQL 결과는 사용자가 원하는 XML문서로 변환하기 위하여 XMLGenerator의 입력으로 제공하고 XMLGenerator는 반환하여야 할 문서의 구조정보와 조합하여 결과 XML문서를 생성한 후 메타데이터 검색기로 반환한다.

3-3. 메타데이터 저장 엔진

TV-Anytime 메타데이터는 XML 문서의 스키마가 미리 정의되어 있다. 즉, 단일 스키마에 근거한 대용량의 메타데이터이다. 그러므로 본 논문에서는 TV-Anytime 메타데이터의 특성에 가장 적절한 데이터베

이스 스키마를 정의하고 있다. TV-Anytime 메타데이터의 구조는 요소(Element)만을 내용으로 가지는 요소가 많다는 특성을 가진다. 그러므로, 본 논문에서는 기본적인 데이터베이스 스키마의 구조는 Edge 단위로 구분하여 동일한 레이블을 가지는 모든 Edge를 하나의 테이블에 저장하는 Binary방법[1]을 사용하지만 요소만을 내용으로 가지는 요소는 상위 부모 노드의 속성으로 취급하여 부모 노드가 저장되는 테이블의 속성으로 저장하는 Inlining방법[4]을 추가함으로써 XQuery에 기술되는 XPath[13]표현을 처리할 때 데이터베이스 테이블간의 불필요한 join의 수가 줄어들도록 하였다. 또한, 본 논문에서는 SQL질의 검색 결과를 다시 XML문서로 재구성 할 경우 보다 효율적으로 문서를 재구성할 수 있고 향후, MPEG-7[5] 데이터를 추가하여 확장할 경우 갱신이 용이 하도록 하기 위하여 Dewey 방법[3]을 적용하여 데이터베이스 스키마를 설계하였다.

메타데이터 기반 방송 시스템에서 방송용 멀티미디어 데이터의 메타데이터를 저장하기 위하여 사용되는 메타데이터 저장 엔진의 구조는 [그림 5]와 같다



[그림 5] 메타데이터 저장 엔진의 구조

TV-Anytime 메타데이터는 대용량이므로 본 연구에서는 대용량 데이터를 효율적으로 저장 관리하기 위하여 상용 관계형 데이터베이스를 사용하였다. TV-Anytime 메타데이터는 XML의 문서 형태를 따르므로 메타데이터 저장 엔진은 TV-Anytime 메타데이터의 저장요청이 오면 해당 XML문서를 구문분석하기 위하여 DOM Parsing[8]하여 DOM Tree를 얻어낸다. Information Extractor에서는 DOM Tree를 순회하면서 메타데이터에 기술되어 있는 내용정보와 구조정보를 얻어낸다. 이렇게 얻어진 내용정보와 구조정보는 InformationLoader를 통하여 TV-Anytime 스키마의 구조를 기반으로 설계된 데이터베이스 테이블에 최적화 하여 저장하게 된다.

4. 결론 및 향후 과제

본 논문에서 제안하고 있는 TV-Anytime 메타데이터 관리 시스템은 메타데이터의 저장을 위하여 데이터베이스를 사용함으로써 대용량 메타데이터를 원활히 관리할 수 있도록 하였으며, XQuery를 메타데이터에 대한 질의로 채택하여 데이터베이스에 독립적인 질의가 가능하도록 하는 동시에 시스템 간의 상호 운용성을 확보하도록 하였다. 현재 프로토타입 시스템을 구현 중에 있다.

급격히 성장하는 XML 기반의 시장에서 XML 데이터의 저장과 검색은 매우 중요한 문제이다. 우리의 연구 결과는 XML 데이터의 저장과 검색이 필요한 다양한 응용에 적용이 가능하며, 특별히 MPEG-7 기반 디지털도서관과 같은 분야에 직접 적용이 가능하다.

구현이 완료되면 본 시스템의 성능을 평가할 예정이다. 앞으로 시스템 성능의 향상을 위하여 데이터베이스 저장 및 XQuery 처리의 최적화를 연구할 필요가 있다. 또한 PDR과 같은 단말기와의 상호 운영을 위하여 단말기에서의 XQuery 처리에 관한 연구 역시 필요할 것이다.

참고 문헌

- [1] D. Florescu & D. Kossmann, "Storing and querying XML data using an RDBMS.", IEEE Data Engineering Bulletin, Volume22, pp 27-34, September 1999.
- [2] I. Manolescu, D. Florescu & D. Kossmann, "Pushing XML Queries inside Relational Databases", INRIA Technical Report, INRIA, No. 4112, January, 2001.
- [3] I. Tatarinov, S.D.Viglas, K.Beyer, J.Shanmugasundaram, E.Shekita, & C.Zhang, "Storing and Querying Ordered XML Using a Relational Database System", Proc. ACM SIGMOD Conf, June 2002.
- [4] J.Shanmugasundaram, K.Tufte, G.He, C.Zhang, D.De Witt, & J.Naughton, "Relational Databases for Querying XML Documents: Limitations and Opportunities", Proc. of 25th VLDB, pp. 302-314, Edinburg, Scotland, September 1999.
- [5] Overview of the MPEG-7 Standard, December, 2001.
- [6] S. Pfeiffer & U. Srinivasan, "TV Anytime as an application scenario for MPEG-7", Proc. ACM Multimedia 2000, Los Angeles, October, 2000.
- [7] TV Anytime Specification Series, August 2001.
- [8] W3C, Document Object Model (DOM) Level 1, Recommendation, Oct. 1998.
- [9] W3C, XQuery 1.0: An XML Query Language, April 2002.
- [10] W3C, Extensible Markup Language (XML) Version 1.0, Recommendation, February 1998.
- [11] W3C, XML Query Language (XQL), September, 1998.
- [12] W3C, XML-QL: A Query Language for XML, August, 1998.
- [13] W3C, XML Path Language (XPath) 2.0, April 2002.
- [14] W3C, XML Schema Part 0: Primer, 2 May 2001.
- [15] MPEG-7 메타데이터 응용 기술, IT Forum korea 2002, May, 2002.
- [16] TV-Anytime Forum 최근동향, IT Forum korea 2002, May, 2002.