

XMF 기반의 가상 멀티미디어 문서 생성*

이강찬**, 이경하, 이규철
충남대학교 컴퓨터공학과
{dolphin, bart, klee}@ce.cnu.ac.kr

Automatic Generation of Virtual Multimedia Documents using XMF

Kangchan Lee, Kyungha Lee, Kyuchul Lee
Department of Computer Engineering, Chungnam National University

요 약

본 논문에서 제안하는 XMF(XML-based Mediation Framework)는 인터넷상의 이질적이고 분산되어 있는 정보 자원에 대하여 사용자에게는 위치 투명성(location transparency)과 접근 투명성(access transparency)을 보장하면서 통합시 발생하는 구조적 충돌(structural conflict)과 의미적 충돌(semantic conflict)을 해결하는 기능을 가지고 있다. 또한 XMF는 통합 모델로써 XML을 채택함으로써 모든 정보 자원에 대하여 모델링 할 수 있게 된다. 본 논문은 이러한 XMF를 기반으로 인터넷에 산재하는 정보 자원, 그리고 DBMS에 저장, 관리되는 데이터를 통합하여 사용자에게 가상 멀티미디어 문서(virtual multimedia document)를 제공하는 시스템을 구축하는 것이다.

1. 서론

인터넷이 보급, 확산된 이후로 인터넷은 하나의 정보 교환 및 검색 수단으로써 자리잡게 되었다[1]. 그러나 인터넷 상의 정보 자원들은 정보 자원의 운영 환경과 구현 언어, 사용하는 데이터 모델, 질의 형식 등에서 서로 많은 이질성을 포함하고 있다[2].

더구나 분산 이질 환경 하에서는 사용자가 정보를 통합 및 공유하기 위해서는 많은 시간과 노력이 필요하게 되며 이는 인터넷상의 정보 자원들이 분산성, 자치성, 이질성, 모호성을 가지고 있기 때문이다

[3].

위와 같은 특성을 갖는 인터넷 정보 자원에 대한 검색은 각 정보 사이트를 일일이 방문하여 정보 자원(information source)으로부터 찾고자 하는 자료를 검색하여야 하며, 동시에 여러 사이트의 정보를 검색하려면 각 사이트에 대하여 개별적으로 검색한 후 그 결과를 사용자 스스로 정리해야 한다. 따라서 다수의 정보 자원으로부터 원하는 것을 검색하여 통합하는 일은 매우 번거로운 작업이라 할 수 있다. 그러나 이러한 통합에 대한 요구는 대규모 인터넷 서비스, 그 중에서도 웹 서비스가 보편화 됨에 따라 증가하고 있으며, 디지털 도서관, CALS/EC, EAI(Enterprise Application Integration)등의 분야에서 그러한 요구를 찾아볼 수 있다. 특히 *가상 문서(virtual document)*, *가상 기업(virtual enterprise)*과

* 본 연구는 충남대학교 부설 소프트웨어공동 연구소(SOREC)의 2000년도 논문연구비 지원을 받았음

**한국전자통신연구원 표준연구센터의 인턴 연구원

같은 경우에는 현재 운영되고 있는 데이터베이스 또는 정보 자원을 특별한 변환작업을 거치지 않고 그대로 사용하며 항상 최신의 데이터를 가지고 새로운 문서, 새로운 데이터를 만들어 내는 정보 통합의 요구가 가장 많이 발생하는 분야이다.

본 논문에서는 가상문서의 정의를 멀티미디어 데이터 통합 문서라 하며 분산 이질 환경에서 DBMS 내의 데이터와 WWW상의 어플리케이션을 통합할 수 있는 XMF[3]를 기반으로 가상문서를 생성 시스템을 구축하기 위한 기본 구조를 제안하며, 통합을 위한 미디어이션 언어(mediation rule)를 기술하는 방법과, XMF의 질의 언어인 Xpath[9]를 이용하여 가상문서에 검색을 하는 방법에 기술하였다.

본 논문의 구성은 다음과 같다. 제2장에서는 가상 문서에 대한 개념을 살펴 본 후에 제3장에서는 XMF에 대한 소개를 한다. 제4장에서는 XMF를 기반으로 가상문서 생성 방법에 대하여 기술하며 제5장에서는 결론 및 향후 연구 과제에 대하여 소개한다.

2. 가상문서

일반적으로 문서는 그 구성요소로서 문서의 내용(content), 문서의 구조(structure), 그리고 문서의 스타일(style)을 가진다. 문서의 내용은 실제 문서의 의미를 나타내는 글자들로 이루어져 있으며 일반적으로 이를 인스턴스(instance)라 부른다. 문서의 구조는 문서의 문서가 가지는 내용을 블록(block)화 시켜 각 블록간의 관계를 나타내는 것으로 이러한 관계는 마크업 언어(markup language)에서는 명시적으로 표현될 수 있으며, 바이너리 형태의 문서에서는 문맥에 의하여 구조를 나타내는 경우도 있다. 문서의 스타일은 문서 내용의 폰트 사이즈, 폰트체, 문단 구조 등과 같은 문서가 프린트 되거나 뷰잉 될 때 사용자에게 문서의 의미를 정확하게 전달하도록 보여지게 만드는 것이다.

가상 문서[4]란 정보 소스로부터 문서 또는 문서의 일부(fragmentation)를 동적(dynamic)으로 만드는 것이다. 가상문서는 동적으로 만들어지기 때문에 문서를 보는 시점에서 그 문서의 내용은 항상 최신

(up-to-date)의 정보만을 가지게 되는 장점이 있다. 가상 문서를 구성하는 문서 내용을 보면, 일반적인 텍스트 이외에 표, 그림 등과 같은 다양한 미디어 타입의 데이터들이 포함될 수 있으며, 이에 가상 문서는 멀티미디어 데이터 통합 문서라 볼 수 있다.

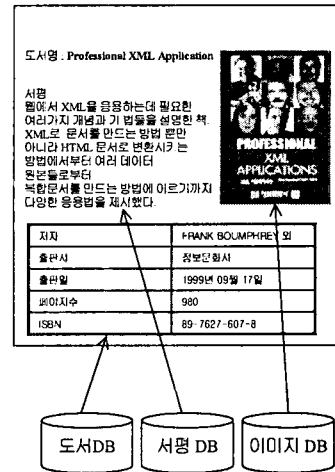


그림 1. 가상 문서의 개요도

3. XMF(XML-based Mediation Framework)

XMF는 XML[5] 기반의 정보 통합 도구로서 미디어이터[6] 구조를 가진다. XMF는 사용자와 래퍼 사이에 위치하면서 중재자 역할을 수행하는 미들웨어(middleware)로서 사용자는 XMF의 미디어이터를 통해 분산된 호스트에 존재하는 정보 자원을 검색할 수 있다. 그림 2는 XMF의 전반적인 구조이다.

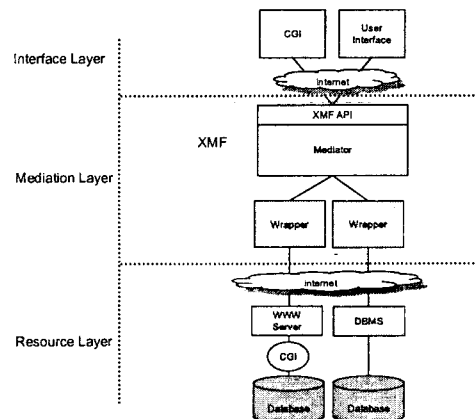


그림 2. XMF의 계층별 구조도

래퍼(wrapper)[7][8]는 미디어이터로부터 받은 질의를 로컬 시스템에 맞게 로컬 질의로 변환하여 수행하고 질의 처리 결과를 미디어이터에 반환하는 프로그램 모듈이다. XMF-ML(XMF Mediation Language)은 Global Schema Definition, Local Schema Definition과 Rule Definition으로 구분되며, 질의 처리나 데이터 통합 시에 참조가 되는 정보들을 기술하는 언어이다. 질의어는 광역 질의 언어(Global Query Language) 사용자가 미디어이터에 질의시 사용하는 질의어-와 서브 질의어 래퍼와 미디어이터 간에 사용되는 질의어 로 나뉘는데 XMF는 두가지 질의어 모두 W3C의 Xpath를 사용한다.

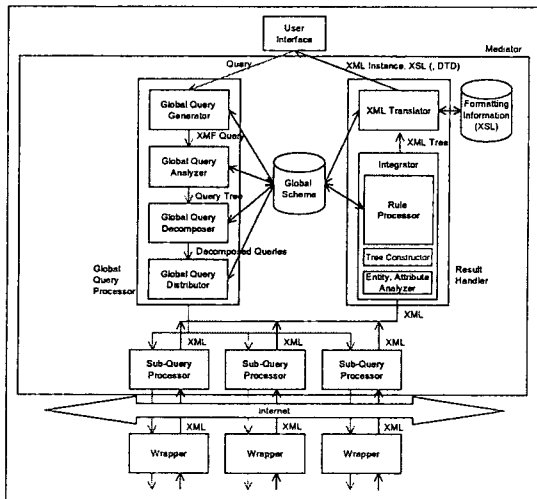


그림 3. XMF 미디어이터 구조

XMF 미디어이터는 사용자와 정보 자원 사이에 위치하면서 사용자로부터 광역 질의를 입력 받아 XMF-ML로 기술된 중재 규칙과 스키마 정보를 이용하여 이를 각 래퍼에 전달할 서브 질의로 분해, 전달 후, 해당 결과를 래퍼로부터 전달 받아 이를 통합하여 사용자에게 전달하는 역할을 수행한다.

4. 가상문서를 위한 XMF

그림 4와 같은 구조의 정보 자원을 통합하여 가상 문서를 만들기 위해서는 각 정보 자원에 대한 로컬 스키마를 작성해야 한다. 로컬 스키마는 각 정

보 자원의 데이터 모델이 서로 상이하기 때문에 같은 데이터 모델을 가지기 위하여 작성하게 된다

로컬 스키마를 작성 한 후에는 글로벌 스키마를 작성하게 되며, 글로벌 스키마는 사용자 입장에서 바라보는 가상 문서의 모습이 된다. 글로벌 스키마 역시 XML 스키마의 형태로 작성된다.

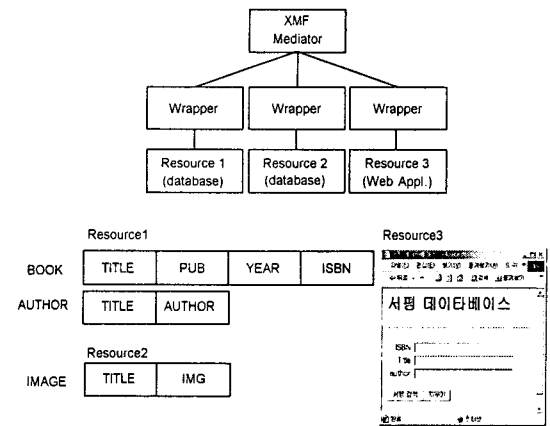


그림 4. 가상 문서를 위한 정보 자원 구조

로컬 스키마와 글로벌 스키마의 작성이 완료되면 이를 기반으로 XMF 중재 언어(mediation language)를 작성한다. XMF 중재 언어는 스키마, 그리고 글로벌 스키마를 구성을 위한 매핑 룰로 이루어진다.

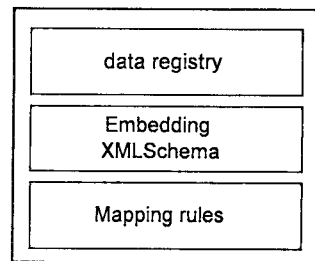


그림 5. XMF 중재 언어의 구조

그림 6은 가상 문서를 위하여 중재 언어를 작성한 일부분으로써 이와 같은 중재 언어를 사용자가 직접 작성하기 어려운 부분이 있으며, 또한 각 엘리먼트와 어트리뷰트의 매핑 관계가 일관(consistency)되게 유지하여야 하기 때문에 매핑을 위한 사용자

인터페이스를 제작중이다.

```
<?xml version="1.0"?>
<DOCTYPE xmf SYSTEM "xmf.dtd">
<xmf xmlns:gs="http://dblab/xmf/book/"
  xmlns:lsl="http://dbleab/xmf/wrapper/lsl/"
  xmlns:lsl2="http://dbleab/xmf/wrapper/lsl2/"
  xmlns:lsl3="http://dbleab/xmf/wrapper/lsl3/">
  <registry>
    <source id="1" name="lsl1" type="DBMS" />
    <source id="2" name="lsl2" type="DBMS" />
    <source id="3" name="lsl3" type="WWW" />
    <dest id="4" name="gs" />
  </registry>
  <schemas>
    <schema id="5" name="lsl1" ref="lsl1.xsd"/>
    <schema id="6" name="lsl2" ref="lsl2.xsd"/>
    <schema id="7" name="lsl3" ref="lsl3.xsd"/>
    <schema id="8" name="gs" ref="gs.xsd"/>
  </schemas>
  <maps>
    <map dest="gs:/book/">
      <map dest="gs:/book/title" source="lsl1:/product/book:title"/>
      <map dest="gs:/book/title" source="lsl2:/results/result/title"/>
      <map dest="gs:/book/title" source="lsl3:/description/title"/>
    </map>
    <map dest="gs:/book:ISBN" source="lsl2:/results/result:isbn"/>
  </maps>
</xmf>
```

그림 6. 가상 문서를 위한 중재 언어

그림 6의 중재 언어에서 글로벌 스키마와 로컬 스키마의 매핑을 나타내는 엘리먼트는 <map> 엘리먼트이다. <map> 엘리먼트의 source 어트리뷰트는 로컬 스키마에 정의된 정보 자원이며 dest 어트리뷰트는 글로벌 스키마에 정의된 정보 자원이다. <gs:/book/title/>과 같이 하나의 글로벌 정보 자원에 대하여 로컬 정보 자원이 여러 개 기술되어 있으면 이는 통합시에 사용되는 키(key)를 의미하며, 만약에 글로벌 정보 자원과 로컬 정보 자원이 1:1로 매칭되어 사용되면 이는 유니온(union) 연산을 의미한다.

중재 언어를 기술한 후에는 XMF 미디어어터를 사용하게 되는데 이때 사용할 XMF 언어는 W3C의 Xpath의 축약형 문법(abbreviate syntax)을 이용한다. Xpath를 질의 언어로 사용하는 이유는 Xpath 자체가 XSL[10] 또는 XSLT[11]에서 사용되기 때문에 이를 이용하여 쉽게 문서 변환에 적용할 수 있는 장점이 있기 때문이다.

5. 결론

XMF는 인터넷상의 이질적인 정보 자원을 XML을 데이터 모델로 통합하며 사용자에게는 단일화된 뷰(view)로 보이도록 하는 통합 시스템이다. 사용자는 XMF를 이용하여 서로 상이한 각각의 정보 자원들에게 일관된 방법으로 질의할 수 있으며, 질의 결과를 통일된 뷰로써 사용자에게 전달한다. 또한 질의 결과는 HTML과 달리 데이터의 구조를 내포하고

있으며, 구조적 질의를 할 수 있다.

본 논문은 위와 같은 XMF의 정보 통합에 대한 기능을 살펴 보고 이를 멀티미디어 데이터를 통합하여 가상 문서를 자동 생성하는데 적용하였다.

향후 연구계획으로는 하이퍼 링크에 대한 관리 및 XMF에 대한 GUI 환경을 구축할 예정이며, W3C의 질의 언어 표준이 완성되면 Xpath를 W3C의 권고안으로 수정할 예정이다.

[참고문헌]

- [1] S.Abiteboul, P. Buneman, D. Suciu, Data On the Web, Morgan Kaufmann Publishers, 2000.
- [2] J.D.Ullman, J. Widom, H. Garcia-Molina, Information Integration, Database System Implementation, Prentice Hall, p.595-612, 1999.
- [3] 이경하, 조정수, 이강찬, 이규철, XMF : XML 기반 분산 이질 정보 자원의 통합 프레임워크, KDBC 2000 학술 발표논문집, p262-270, 2000.
- [4] Anne-Marie Vercoustre, Francois Paradis, A Descriptive Language for Information Object Reuse through Virtual Documents, In proceeding of International Conference on Object Oriented Information Systems, pp. 299-311, 1997.
- [5] Tim Bray and C.M. Sperberg-McQueen, "Extensible Markup Language (XML): Part I. Syntax", World Wide Web Consortium Recommendations, February 1998, Available at <http://www.w3.org/TR/REC-xml>.
- [6] Gio Wiederhold, Mediators in the Architecture of Future Information Systems, The IEEE Computer Magazine, 25(3):38-49, March 1992.
- [7] 김경일, 이경하, 이강찬, 이규철, XML RDB 계이트웨이의 설계 및 구현, 한국정보과학회 2000 가을 학술발표논문집, 2000.
- [8] 김태현, 김경일, 이강찬, 이규철, XML 기반의 정보 통합을 위한 OODB2XML 래퍼의 설계 및 구현, 한국정보과학회 99 가을 학술발표논문집(1), 제 26권 제 2호, pp. 75-77, 1999.
- [9] James Clark, Steve DeRose, "XML Path Language (XPath) Version 1.0", World Wide Web Consortium Recommendations, Nov. 1999, Available at <http://www.w3.org/TR/xpath>.
- [10] James Clark, and Stephen Deach, "Extensible Stylesheet Language (XSL) Specification", April 1999, Available at <http://www.w3.org/TR/WD-xsl>.
- [11] James Clark, "XSL Transformations (XSLT) Version 1.0", November 1999, Available at <http://www.w3.org/TR/xslt>.