

XML기반의 데이터 저장관리를 위한 RDB2XML 변환 레퍼

장우혁, 김홍식
인제대학교 전산학과

RDB2XML Wrapper for XML-based Data Storage Management

Woo-Hyuk Chang^o, Heung-Shik Kim
Dept. of Computer Science, Inje University

요 약

W3C에서 제안한 XML 표준안이 발표된 이후 인터넷 기반의 다양한 분야에서 XML 기술을 활용하고자 하는 움직임이 많아지고 있다. 그 중 한 분야가 기존의 데이터베이스 구조를 이질의 플랫폼에 영향을 받지 않고 XML을 매개로 하여 저장 및 관리가 가능한 시스템을 구축하고자 하는 연구이다. 이는 분산 환경에서의 각자의 고유한 데이터베이스 스키마를 가지고 운영되는 기존의 정보 시스템들을 XML이라는 공통 데이터 모델을 사용하여 표현함으로써 저장 및 관리상에 효율성 증대를 제공하는 것이다

1. 서론

B2B(Business to Business)시대에는 XML문서의 양도 많아지며, 이것을 저장,관리할 XML저장 관리 기술이 필요하다. 또한 데이터베이스 관련 아키텍처를 XML문서로 변환하여 저장하는 기술도 필요하다. 이러한 데이터베이스 관련 아키텍처를 데이터 형태로 저장하기 위해서는 XML문서와 데이터베이스간에 매핑이 필요하다. XML 문서의 구조적 정보를 토대로 데이터베이스에 그 형태대로 저장하는 방식으로 여러 가지 알고리즘에 의해 구현이 가능하다[3].

또한, 이질의 플랫폼에서 데이터베이스간에 호환성과 효율성을 기하기 위해서 데이터베이스와 XML간의 매핑을 정의하고, 이러한 매핑을 통해서 데이터베이스 스키마를 XML문서로 변환하고, 변환된 XML 문서를 다시 데이터베이스 스키마 구조로 복원하고, XML문서를 통해서 기존의 데이터베이스 관리 시스템처럼 데이터의 조작 및

정의, 검색이 가능한 Wrapper의 설계가 필요하다. 이러한 연구에서는 기존의 데이터베이스 관리 시스템의 스키마 구조를 XML로 변환하는 방법과 변환된 XML 파일에서 기존의 데이터베이스 시스템처럼 조작 및 정의 기능이 가능하도록 하는 XML DTD에 유효한 정의가 필요하다.

따라서, 본 논문에서는 이질의 플랫폼에서도 기존의 관계형 데이터베이스를 XML 파일로 변환, 복원이 가능한 시스템을 설계하고자 한다.

본 논문은 다음과 같이 구성되었다. 제2장에서는 정보교환의 표준으로 등장한 XML과 XML 저장 기술에 대해서 알아보고, 자바와 XML에서 데이터베이스 연동을 위한 JDBC에 대해서 알아본다. 제3장에서는 DB2XML 변환 Wrapper의 각각의 구성요소의 기능 및 설계에 대해서 알아보고, 제4장에서는 설계된 DB2XML 변환 Wrapper를 구현해 보고자 한다. 마지막으로 제5장에서는 DB2XML 변환 Wrapper에 대한 향후 개선책 및

결론을 기술한다.

2. XML 저장기술

수 많은 기업들이 B2B 포털 사이트 구축을 XML로 검토, 진행중인 상황에서 대용량의 XML 문서를 저장, 검색, 관리하기 위해서는 XML 특성에 맞는 저장 관리 시스템의 구축이 필요하다. 이러한 저장 관리 시스템 구축을 위해서는 다음과 같은 XML 문서 저장시의 고려사항을 염두해 두어야 한다[1][2][4].

- XML 모델의 무결성(XML Model Integrity)
저장 시스템에 저장하기전의 XML 문서와 저장후 다시 복원한 XML 문서는 완전히 일치해야 한다.
- 대용량의 데이터(Large Data)
멀티미디어와 같은 데이터를 포함한 대용량의 XML문서를 다룰 수 있어야 한다.
- 메타데이터(Metadata)
어떠한 형태의 메타데이터도 요소가 될 수 있어야 한다.
- 엔티티 관리(Entity Management)
모든 데이터 타입에 대해 인덱스 될 수 있어야 한다. 즉, Entity에 대한 관리를 충분히 해야 한다.
- 계층적 네임스페이스(Hierarchical Namespace)
계층적 네임스페이스를 충분히 지원해서 XML문서 내용이 데이터베이스와 데이터베이스간의 분리되어 있어도 사용 가능하도록 하여야 한다.

2.1.1 XML 저장관리방식

① 저장방식에 따른 분류

현재 HTML을 저장하는 형태로써 XML 문서를 유지하며 저장하는 방법으로 파일 시스템에 저장하거나 데이터베이스 시스템에 저장하는 방법이 있으며, 이는 문서를 저장할 때 유용하다. 만일 XML 문서가 잘 정의된 구조이고, 데이터중심의 문서일 경우 문서를 데이터 형태로 변환하여 저장이 가능하다.

② 데이터베이스 방식에 따른 분류

데이터베이스 방식에 따른 분류에는 RDBMS와 OODBMS에 저장하는 방식으로 나누어 설명할 수 있다. RDBMS에 저장하는 방식은 현재 RDBMS 시장이 다른 DBMS에 비해 넓고

사용자 수도 그에 따라 많기 때문에 확산이 빠르며 쉽게 상용화할 수 있다는 장점이 있다. 일반적인 XML 파일이 데이터 형태로 변환되어 데이터베이스에 저장하는 과정을 그림1에서 나타내고 있다.

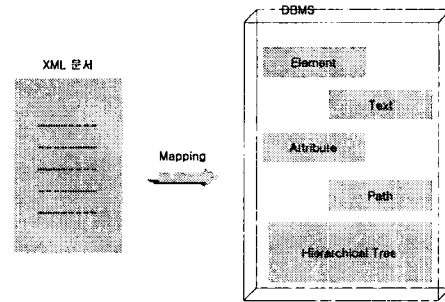


그림 1 데이터 변환 저장

3. DB2XML 변환 Wrapper의 설계

3.1 DB2XML 변환 Wrapper 의 개요

본 논문에서 설계 및 구현하고자 하는 DB2XML 변환 Wrapper는 기존의 데이터베이스 관리 시스템의 스키마 구조를 이질의 플랫폼과 관계없이 XML파일로 변환하여 변환된 XML파일을 데이터베이스화하여 데이터의 조작과 정의, 검색 기능을 할수 있도록 한다. 또한 변환된 XML파일을 다시 기존의 데이터베이스 구조로 복원할 수 있도록 하여 데이터베이스와 XML파일 간에 상호 변환이 가능하도록 하였다.그림2에서는 DB2XML 변환 Wrapper 구조를 나타내고 있다.

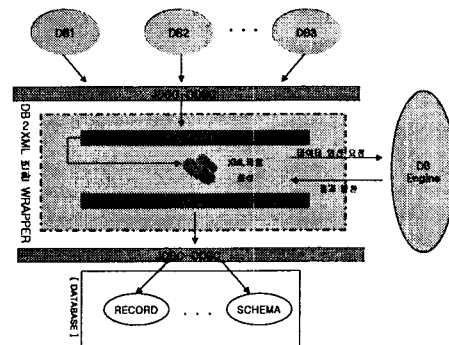


그림 2 DB2XML 변환 Wrapper 구조

3.2 DB2XML 변환기

DB2XML 변환 Wrapper를 구성하는 변환기들은 W3C에 제안된 DOM 빌더와 SAX 파서를 이용하여

구현된다. DB2XML 변환기는 DOM 빌더를 이용하여 데이터베이스로부터 XML문서를 만들어내고 SAX파서를 이용하여 XML파일을 해석해서 다시 데이터베이스에 레코드값을 입력하게 된다. DB2XML 변환 Wrapper에서 XML파일로 변환하고자 하는 데이터베이스는 관계형 데이터베이스이다. 이는 아직까지 B2B 기반의 데이터베이스 시스템을 구축하고자하는 기업들의 약 70~80%가 관계형 데이터베이스를 사용하고 있기 때문이다. 그림3은 DB2XML 변환 Wrapper에서 XML파일을 처리하는 과정을 나타내고 있다.

```
<?xml version="1.0" encoding="euc-kr" ?>
<!-- Automatically generated by the db2xml conversion tool -->
<DOCTYPE database [
<ELEMENT database (records)*>
<ELEMENT records (record)*>
<ELEMENT record (memberkey, attenddate, graddate, schoolname, major)*>
</ELEMENT record >
</ELEMENT records >
</ELEMENT database >
</DOCTYPE >
<database>
<records>
<record>
<memberkey>
1
</memberkey>
<attenddate>
1980-03-01 00:00:00
</attenddate>
<graddate>
1987-03-01 00:00:00
</graddate>
<schoolname>
seoul
</schoolname>
<major>
computer
</major>
</record>
<record>
<memberkey>
2
</memberkey>
<attenddate>
1975-02-28 00:00:00
</attenddate>
<graddate>
1980-03-03 00:00:00
</graddate>
<schoolname>
busan
</schoolname>
</records>
</database>
```

그림3 데이터베이스 스키마 구조를 XML 파일로 변환

3.3 XML2DB 변환기

XML2DB 변환기는 데이터베이스 스키마 정보를 복원하는 작업을 한다. 이는 XML 문서를 파싱하는 과정에서 그림4에서 만들어진 엘리먼트 트리의 엘리먼트 이름과 속성을 추출하여 스키마 정보로 복원하는 과정을 거치게 된다. 이러한 추출 과정을 위해서 SAX파서의 startElement 메소드에 엘리먼트 이름과 엘리먼트 속성이 넘어오게 되므로, RECORD 엘리먼트가 시작되는 시점에서 엘리먼트에 속한 하부 엘리먼트에 대한 정보를 추출하여 스키마 정보를 손쉽게 복원할 수 있다. 그림4는 XML파일에서 데이터베이스의 스키마 정보를 추출하여 저장하는 방식을 나타내고 있다.

4.2 XML2DB 변환기 구현

XML2DB 변환기는 앞서 변환된 XML 파일을 원래의 데이터베이스 스키마 구조로 복원하는 역할을 한다. 그림4는 앞서 변환된 XML파일을 기존의 데이터베이스 스키마 구조로 복원된 모습을 나타낸다.

3.4 DB Engine 설계

DB Engine은 XML파일을 기존의 관계형 데이터베이스에서 연산을 가능하도록 만들어주는 기능을 한다. 따라서, 연산기능으로 삽입, 삭제, 검색, 갱신 등의 SQL 전반적인 연산기능을 제공한다. 그림5은 관계형 데이터베이스를 XML DTD파일로 매핑하는 과정을 나타내고 있다. 우선 관계형 데이터베이스의 스키마를 테이블과 필드로 나누고 필드의 일부를 속성값으로 매핑시킨다. 이렇게 매핑된 XML 파일을 이용해서 데이터베이스 연산을 수행하게 되는 것이다

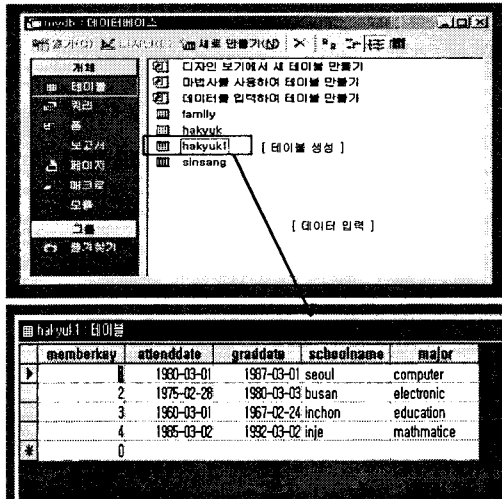


그림4 데이터베이스 데이터 타입에 대한 XML DTD의 정의

4. DB2XML 변환 Wrapper의 구현

4.1 DB2XML 변환기 구현

관계형 데이터베이스 스키마 구조와 인스턴스를 이용하기 위해 먼저 ODBC를 설정하고 DOM 빌더를 이용해서 데이터베이스 스키마 구조를 XML 파일로 변환하면 그림3과 같은 결과가 나타난다.

4.3 DB Engine 구현

앞서 XML2DB 변환기에 의해 생성된 XML파일을 이

용하여 데이터베이스의 조작 기능을 제공하는 DB Engine은 기존의 데이터베이스의 대부분의 조작 기능을 제공해 준다. DB Engine에서 제공하는 조작 기능으로 삽입, 삭제, 추가, 변경, 전체 레코드 및 엘리먼트 리스트 출력, 검색등의 다양한 기능을 제공해 준다. 그림5는 DB Engine의 초기화면을 나타내고 있다. 초기 화면에서 다양한 데이터의 조작이 가능하다

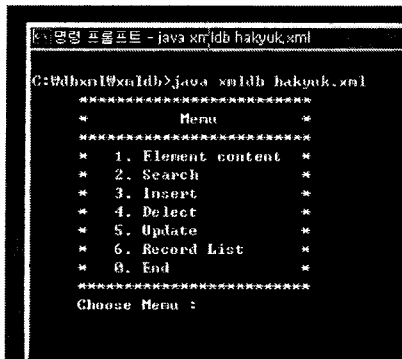


그림5 DB Engine의 초기 화면

5. 결론 및 향후 계획

본 논문에서는 XML파일을 이용하여 기존의 관계형 데이터베이스간의 상호 변환과 조작이 가능한 Wrapper를 설계하였다. 이질의 데이터나 플랫폼에 관계없이 XML파일과 데이터베이스간의 상호 변환이 용이하도록 하였다. 또한 변환된 XML파일상에서 삽입, 삭제, 갱신, 검색 등의 조작기능을 가능하게 하였다. 향후 연구 방향으로는 기존의 관계형 데이터베이스뿐만 아니라 객체지향 데이터베이스간의 상호변환이 가능한 시스템의 설계가 필요하며, 또한 웹 상에서 사용자의 정보에 대한 요청을 XML파일로 변환하여 쉽고 표준화된 검색이 가능하도록 하는 시스템의 구현이 필요하다.

6. 참고 문헌

[1] Charles F.Goldfarb and Paul Prescod, The XML Handbook Prentice Hall, 2000
 [2] Jennifer Widom, Data Management for XML, White Paper
 [3] Charles Phillips and Mary Meeker, The B2B Internet Report, Morgan Stanley Dean Witter, 2000
 [4] XMLDatabaseProduct - <http://www.xmlsoftware.com/database>

[5] Extensible Markup Language(XML) 1.0 <http://www.w3.org/TR/PR-xml-971208>
 [6] Document Object Model(DOM) Level 1 Specification <http://www.w3.org/TR/1998/PR-DOM-level-1-19980818>
 [7] W3CDTD : <http://www.w3.org/TR/2000/REC-xml-20001006>
 [8] Eva Maler, Jeanne El Andaloussi, "Developing SGML DTDs From Text To Model To Markup", Prentice-Hall, 1996
 [9] Didier Martin and Mark Birbeck and Michael Kay and Brian Loesgen and Jon Pinnock and Steven Livingstone and Peter Stark and Kevin Williams and Richard Anderson and Stephen Mohr and David Baliles and Bruce Peatand Nikola Ozu, Professional XML, WROX press, 2000
 [10] Martin Bryan, An Introduction to the Extensible Markup Language(XML), <http://www.personal.u-net.com/~sgml/xmlintro.htm>
 [11] 이용석, 손기락, "XML 문서 저장 시스템 설계 및 구현", 정보과학회 학술 발표 논문집(I), 25권 2호, 1998
 [12] Serge Abiteboul, Peter Buneman, Dan Suciu, Data on the Web From Relations to Semistructured Data and XML, Morgan Kaufmann Publishers, 2000
 [13] Ronald Bourret, XML and Databases <http://www.rpbouret.com/xml/XML And Databases.htm#isxmladatabase>
 [14] Ken Arnold, James Gosling, "The Java TM Programming Language," Addison-Wesley, 1996
 [15] David Flanagan, "Java In A Nutshell : A Desktop Quick Reference 2nd Edition," 1997 O'Reilly
 [16] Hiroshi Maruyama, Kent Tamura, Naohiko Uramoto, "XML and Java Developing Web Applications," Addison-Wesley, 1999