

u-GIS 환경에서 효율적인 공간 정보 유통을 위한 S-XML 변환 기법

(S-XML Transformation Method for Efficient Distribution
of Spatial Information on u-GIS Environment)

이 동욱* 백 성 하* 김 경 배** 배 해 영***
(Dong Wook Lee) (Sung Ha Baek) (Gyoung Bae Kim) (Hae Young Bae)

요약 u-GIS 환경에서는 센서 네트워크를 통해 필요한 공간 데이터를 수집하고 이를 실시간 처리 및 가공 또는 기 저장 되어 있는 정보와 함께 유통된다. 웹 기반 응용서비스 등에서 인터넷 망을 통한 정보가 요청되는 경우 표준 문서인 XML로 전달된다. 특히 요청되는 정보에 공간 데이터가 포함되는 경우 공간데이터 처리가 가능한 GML, S-XML 등의 문서가 사용된다. 이 과정에서 DSMS에서와 같이 실시간 처리된 스트림데이터는 S-XML 문서 형태로 변환되고, 웹 기반의 공간정보 응용서비스는 인터넷 망을 통해 S-XML 문서를 전달 받는다. 대부분의 공간정보 응용서비스는 저장시스템으로 기존의 공간 데이터베이스 관리 시스템을 사용하기 때문에 S-XML 데이터와 SDBMS에서 사용되는 데이터간의 상호 변환과정이 필요하다. 본 논문에서는 공간 데이터의 캐싱을 이용한 S-XML 변환 기법을 제안한다. 제안 기법은 공간 정보 유통을 위한 S-XML과 관계형 공간 데이터베이스와의 효율적인 변환을 위해, S-XML에서 공간 데이터에 해당하는 부분을 캐싱하고, 동일 지역의 공간데이터에 대한 변환이 요구될 경우 캐시 데이터를 재사용하여 별도의 변환 비용 없이 변환한다. 제안 기법을 통해 u-GIS 환경에서 공간정보의 유통을 위한 S-XML 문서와 이를 이용하는 웹 기반 공간정보 응용서비스 사이의 변환 비용을 감소하였으며, 성능평가를 통하여 질의 처리 성능이 향상됨을 보인다.

키워드 : 공간정보 유통, S-XML 변환, 공간데이터 캐싱

Abstract In u-GIS environment, we collect spatial data needed through sensor network and provide them with information real-time processed or stored. When information through Internet is requested on Web based applications, it is transmitted in XML. Especially, when requested information includes spatial data, GML, S-XML, and other document that can process spatial data are used. In this processing, real-time stream data processed in DSMS is transformed to S-XML document type and spatial information service based on web receive S-XML document through Internet. Because most of spatial application service use existing spatial DBMS as a storage system, The data used in S-XML and SDBMS needs transformation between themselves. In this paper, we propose S-XML a transformation method using caching of spatial data. The proposed method caches the spatial data part of S-XML to transform S-XML and relational spatial database for providing spatial data efficiently and it transforms cached data without additional transformation cost when a transformation between data in the same region is required. Through proposed method, we show that it reduced the cost of transformation between S-XML documents and spatial information services based on web to provide spatial information in u-GIS environment and increased the performance of query processing through performance evaluation.

Keywords : Spatial Information Distribution, S-XML Transformation, Spatial Data Caching

1. 서론

u-GIS는 유비쿼터스 환경을 기반으로 공간정보 기술을 활용하는 것으로 시공간의 제약을 받지 않고 사용자별 맞

춤형 공간정보 서비스를 제공하는 것이다[1]. 이를 위해 RFID리더, 센서 노드, 텔레메틱스 단말, CCTV 등으로 구성된 GeoSensor로부터 온도, 습도, 위치 등의 공간 데이터를 실시간 스트림 형태로 획득한다.

[†] 본 연구는 건설교통부 첨단도시기술개발사업- 지능형국토정보기술혁신 사업과제의 연구비지원(07국토정보C05)에 의해 수행되었음.

* 인하대학교 정보공학과 박사과정, dwlee@dlab.inha.ac.kr(교신저자), shbaek@dlab.inha.ac.kr

** 서원대학교 컴퓨터교육학과 조교수, gbkim@seowon.ac.kr

*** 인하대학교 정보공학부 교수, hybae@inha.ac.kr

XML은 현재 웹 환경에서 데이터 유통의 표준으로 사용되고 있으며, 웹 응용서비스를 위한 XML 전용 DBMS가 개발되는 등 활용도가 높아지고 있다[2, 3]. OGC (Open GIS Consortium)에서는 XML에서 공간 데이터 표현을 위한 GML(Geography Markup Language)을 제안하여 공간 데이터의 웹 유통을 용이하게 하고 있다[4, 5]. 또한 모바일 환경을 위한 G-XML, 음성 스키마를 지원하는 S-XML(Spatial-eXtensible Markup Language)과 같은 확장된 형태의 문서들이 제안되었다[6, 7].

인터넷 망 기반의 u-GIS 공간 응용서비스는 실시간 맞춤형 서비스를 위해 실시간 공간, 비공간 데이터와 구축된 다양한 공간정보를 요구한다. 실시간 데이터스트림 정보는 GeoSensor 및 센서 네트워크에서 생성된 실시간 데이터스트림을 DSMS에서 사용자의 연속질의 등으로 처리하여, 이를 GML, S-XML 등의 공간 정보가 포함된 XML 문서 형태로 인터넷 망을 통해 전달한다. 또한 대부분의 공간정보 응용서비스는 저장시스템으로 기존의 공간 데이터베이스 관리 시스템(SDBMS: Spatial Database Management System)을 사용하기 때문에 필요한 공간정보를 S-XML 형태로 전달 받은 경우 XML 데이터와 SDBMS에서 사용되는 데이터간의 상호 변환과정이 필요하다. 그림 1은 웹 기반의 공간 응용서비스를 위한 S-XML 유통 흐름을 나타낸다.

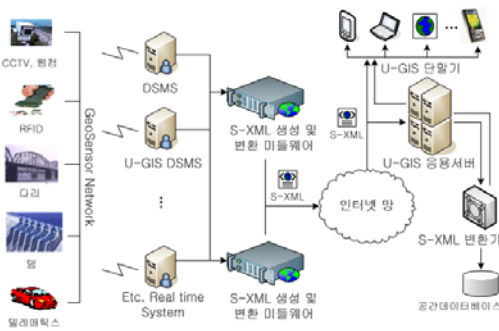


그림 1. u-GIS환경에서 공간 응용서비스를 위한 S-XML 유통 흐름도

본 논문에서는 공간 데이터의 캐싱을 이용한 S-XML 변환 기법을 제안한다. 제안 기법은 공간 정보 유통을 위한 S-XML과 관계형 공간 데이터베이스와의 효율적인 변환을 위하여 필요한 공간정보가 S-XML로 전달되어 변환과정이 수행될 경우 S-XML에서 공간 데이터 부분을 캐싱하여, 캐시 데이터에 대한 색인을 생성하여 관리한다[8]. 이후 검색 혹은 동일 지역의 공간데이터에 대한 변환이 요구될 경우 캐시 데이터를 재사용하여 별도의 변환 비용 없이 변환한다.

본 논문의 구성은 다음과 같다. 2장에서는 관련연구 내용으로 공간 데이터 유통을 위한 기존의 XML 변환 기법들을 살펴본다. 3장에서는 S-XML 관리 시스템에서 확장된 S-XML 변환기의 구조와 효율적인 변환을 위한 캐싱

과정과 변환과정을 기술한다. 4장은 기존의 S-XML 변환기와 확장된 S-XML 변환기에서의 변환 성능을 비교 평가하고, 5장에서 결론 및 향후 연구과제에 대해 논의한다.

2. 관련 연구

관계형 DBMS를 저장 시스템으로 사용하는 공간 응용 서비스를 위한 기존의 XML 변환 방법 및 S-XML 관리 시스템을 설명한다.

2.1 스키마 컨버전을 통한 XML 변환 기법

XML 문서와 관계형 데이터베이스 시스템의 테이블 간의 효율적인 스키마 컨버전을 통한 변환 기법은 관계형 데이터베이스의 테이블 구조를 계층적인 XML 구조로 변환하기 위한 알고리즘과 스키마를 정의하여, 이를 기반으로 동작하는 XML 문서 구축기를 사용하였다[9]. 이는 RDB의 처리 결과를 XML 문서의 계층적인 구조로 변환하기 위한 연산과, 여러 개의 테이블이 관계를 이루는 형태일 때 XML 구조를 정의하는 연산을 정의하였다. 또한 연산 결과를 최종적으로 XML 문서에 반영하기 위한 스키마 디자인어를 사용하였다.

본 관련연구에서는 DBMS에서 처리된 결과에서 XML로의 변환 연산비용 감소 위한 기법이다. 그러나 공간 데이터의 경우 데이터 자체의 크기와 복잡도가 크기 때문에 변환 연산비용 감소가 쉽지 않다. 또한 변환 비용이 큰 공간데이터의 경우 변환 횟수를 줄이는 방법이 모색되어야 한다.

2.2 캐시를 이용한 XML 질의처리 기법

XML 지원 시스템에서 질의처리 성능을 향상하기 위해 XML 구조의 실체화 뷰를 유지하는 연구가 진행되었다[10, 11]. 이 연구는 XQuery에서 정의되는 XML 실체화 뷰를 이용하여 질의처리 성능을 향상하였고, 자료구조로 MIT(Modified Incomplete Tree)를 사용한다[12, 13].

다음으로 XQuery의 구문에서 사상과 변환과정이 이루어지는 알고리즘을 제시하였다. 질의 및 실체화 뷰를 정의하는 두 개의 XQuery에서 패턴 변수를 사용하여 질의 포함 사상과 질의 변환을 수행하는 알고리즘으로, 이를 이용한 ACE-XQ라는 프로토타입 시스템을 개발하였다.

XML 질의처리를 위하여 기존의 캐싱을 이용한 연구에서는 XML 지원을 위한 시스템의 질의처리 비용을 줄이고 부하 분산을 위해 DBMS와 연동되는 데이터 서버 또는 응용 서버에 XML 질의 결과를 캐싱하거나 실체화 뷰를 생성하여 이용하였다[14].

위의 관련연구에서는 XML 질의의 결과 XML 데이터를 데이터 서버 또는 응용 서버에 캐싱하여 질의처리 시 캐싱되어 있는 데이터를 이용하고, 캐시 데이터에 존재하지 않는 데이터는 DBMS에서 처리되어 기존의 변환과정이 필요하다. 따라서 변환과정 자체의 부하는 고려되고 있지 않고, 특히 공간데이터를 저장하는 SDBMS에서는 공간 데이터의 변환 비용이 질의처리 성능에 영향을 미친다.

2.3 S-XML 관리 시스템

S-XML 관리 시스템은 S-XML의 Simple Feature 명세를 기반으로 상호 이질적인 공간 데이터 유통을 위한 시스템으로 S-XQuery를 이용하여 공간 정보를 추출한다 [15]. 그림 2는 S-XML 관리 시스템의 구조로 S-XQuery 질의 변환기, 응용 스키마 생성기, S-XML 변환기, 다큐먼트 저장소, S-XML 분석기로 구성된다.

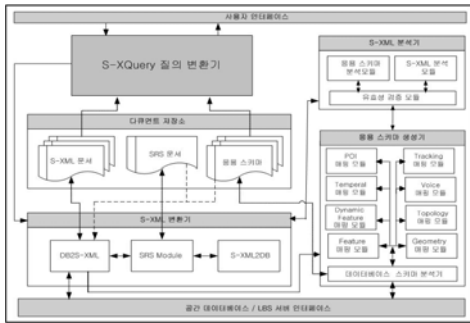


그림 2. S-XML 관리 시스템 구조

S-XML 관리 시스템은 SDBMS에서 처리된 이진 데이터를 S-XML로 변환하는 DB2S-XML과 S-XML 문서를 SDBMS로 저장하기 위한 S-XML2DB 모듈이 주요 변환 작업을 수행한다. 하지만 비용이 큰 공간 데이터에 대한 변환 작업이 S-XML 정보의 입출력 시 반복되어 시스템 성능 및 이를 기반 한 공간 응용서비스들에 병목 현상의 요인이 된다.

3. 공간 데이터 캐싱을 이용한 S-XML 변환

이 장에서는 S-XML 관리 시스템의 S-XML 변환기에서 변환성을 향상을 위하여 확장된 변환의 구조 및 자료구조, 캐시 구조와 질의처리 방법에 대해 설명한다.

3.1 효율적인 S-XML 변환을 위한 S-XML 변환기

본 장에서는 확장된 S-XML 변환기의 구조에 대해서 설명한다. S-XML 변환기는 DB2S-XML 모듈에 캐시 데이터를 위한 캐시 관리자(Cache Manager)와 S-XQuery 분석기(S-XQuery Analyzer)로 이루어진다. 그림 3은 S-XML 변환기의 구조이다.

S-XQuery 분석기는 입력 된 S-XML 문서의 스키마를 분석하고, 문서의 유효성을 검증한다. S-XML 문서에 포함된 질의가 입력되면 문서의 유효성을 검증하기 위해 데이터베이스의 데이터 스키마를 참조하여 처리여부를 결정한다.

S-XML 변환기는 사용자로부터 S-XML 문서 형태로 질의를 입력 받고, SDBMS에서 처리한 이진 데이터를 S-XML로 변환하여 사용자에게 전달한다. 변환된 질의는 SDBMS 인터페이스를 통하여 SDBMS에 전송되고,

SDBMS는 처리된 결과를 해당 지원 시스템에 전달한다.

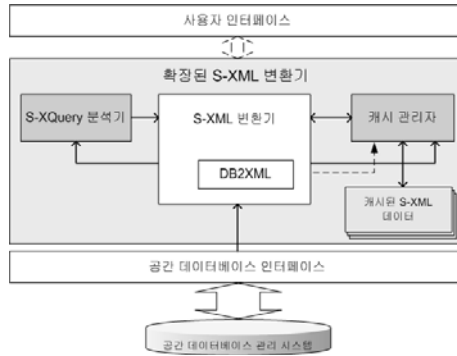


그림 3. 확장된 S-XML 변환기 구조

확장된 S-XML 변환기에서는 SDBMS에서 처리한 이진 데이터를 S-XML로 변환하기 위해, S-XQuery 분석기를 이용하여 질의 결과에 대한 분석과정이 필요하다. S-XQuery의 파싱된 정보를 참조하여 사용자가 추가로 정의의 엘리먼트 데이터를 추출하여 DOM 트리를 이용하여 S-XML 문서로 변환한다.

S-XML 변환기에서 SDBMS의 이진 데이터의 변환 결과로 S-XML 문서가 생성되면 캐시 관리자는 질의형태에 따라 사용할 색인을 결정하여, S-XML에서 공간 데이터에 해당되는 데이터를 캐싱한다.

캐시 관리자는 변환 컴포넌트의 출력 데이터인 S-XML의 데이터 중에서 공간 데이터에 해당하는 부분만을 캐싱하여 메인 메모리에 저장하고 S-XML 변환기에서 요청하는 데이터 중 캐싱되어 있는 데이터를 검색하여 제공한다. 캐시 데이터는 S-XML 데이터의 형식을 보존하고 질의 분석기의 질의 종류에 따라 점 기반 질의와 영역 기반 질의로 구분하여 처리한다.

캐싱 오퍼레이터는 S-XML 분석기로부터 질의 결과에 대한 정보를 참조하여 추출 명령을 계획하고 결과 데이터에 적합한 색인을 사용하여 추출된 공간 데이터를 저장 및 관리한다. 또한, 질의 형태에 맞는 색인을 사용하여 캐시 데이터를 검색하여 S-XML 변환기에 제공한다. S-XML 문서가 점 질의에 의한 결과 문서일 경우 캐싱한 공간 데이터의 OID값을 키로 하는 해시를 사용하여 저장하고, 영역 질의의 경우에는 R-tree를 이용하여 저장한다. 공간 데이터 추출기는 캐싱 오퍼레이터의 명령에 따라 S-XML에서 공간 데이터를 검색한 뒤, 추출하여 캐싱 오퍼레이터에 전달한다.

캐시 데이터는 메인 메모리에서 저장 및 관리되고, 공간데이터의 계층별로 캐시 데이터를 구축한다. 예를 들어 공간 데이터가 '강', '도로', '건물' 등의 계층 구조로 이루어져 있다면, '강'에 대한 공간 데이터를 위해 해싱과 R-tree가 사용되고, '도로', '건물' 등도 같은 방법으로 별도의 캐시 데이터로 저장된다. 그림 4는 캐시 관리자의 구조를 나타낸다.

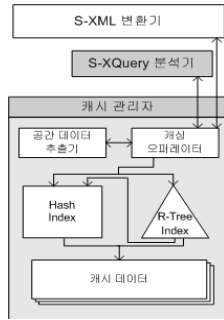


그림 4. 캐시 관리자 구조

3.2 공간 데이터 캐싱을 위한 자료구조

제안 기법은 S-XML과 SDBMS 사이의 효율적인 상호 변환을 위해 S-XML에서 공간데이터에 해당하는 부분만을 추출하여 캐시 관리자에 의해 캐싱하고, 변환과정에서 이용한다. 본 장에서는 S-XML 문서 내에 존재하는 공간 데이터를 캐싱하기 위한 자료구조를 설명한다.

S-XML 문서를 변환하는 과정에서, 캐시 관리자는 S-XML의 공간 데이터를 ‘<OBJ>’ 태그가 포함하는 데이터를 캐싱하고 <OID> 태그의 OID 값으로 색인을 구축한다.

```
<?xml version="1.0" encoding="EUCKR" ?>
<?xml-stylesheet href="global.xml" type="text/xsl" ?>
<BuildingSchema xmlns:gmi="http://www.opengis.net/gmi" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="file:///K:/wwwroot/gmi/Building.xsd">
<gmi:boundedBy>
<gmi:Box>
<gmi:coord><gmi:X>196575</gmi:X><gmi:Y>435627</gmi:Y></gmi:coord>
<gmi:coord><gmi:X>208821</gmi:X><gmi:Y>448112</gmi:Y></gmi:coord>
</gmi:Box>
</gmi:boundedBy>
<BuildingSchemaMember>
<Building>
<OBJ>
<gmi:Polygon>
<gmi:outerBoundaryIs>
<gmi:LinearRing>
<gmi:coord><gmi:X>199047</gmi:X><gmi:Y>438327</gmi:Y></gmi:coord>
<gmi:coord><gmi:X>199052</gmi:X><gmi:Y>438322</gmi:Y></gmi:coord>
<gmi:coord><gmi:X>199056</gmi:X><gmi:Y>438327</gmi:Y></gmi:coord>
<gmi:coord><gmi:X>199052</gmi:X><gmi:Y>438331</gmi:Y></gmi:coord>
<gmi:coord><gmi:X>199047</gmi:X><gmi:Y>438327</gmi:Y></gmi:coord>
</gmi:LinearRing>
<gmi:outerBoundaryIs>
<gmi:Polygon>
<OBJ>
<OID>1</OID>
</ID>1</ID>
</Building>
</BuildingSchemaMember>
```

그림 5. 공간 데이터를 포함한 S-XML 문서 예

캐시 관리자는 그림 6과 같은 자료구조를 사용하여 S-XML에서 공간 데이터를 캐싱하여, 캐시 데이터를 구축한다. S-XML에 포함된 공간 데이터가 추출되면, 첫 번째로 영역 질의처리에 대한 공간 데이터를 캐시 데이터에서 검색하기 위하여 R-tree를 이용하여 공간 데이터의 OID를 삽입한다.

OID	OID Number
Transformation Data	<OBJ>...</OBJ>

그림 6. 공간 데이터 캐싱을 위한 데이터 구조

R-tree의 루트 노드와 중간 노드의 구성은 일반적인 공간 데이터를 다루는 R-tree와 같지만, 리프 노드에는 캐싱된 데이터의 주소 값과 이를 참조하는 OID가 저장된다. 그림 7은 R-tree로 구성된 캐시 데이터의 구조이다.

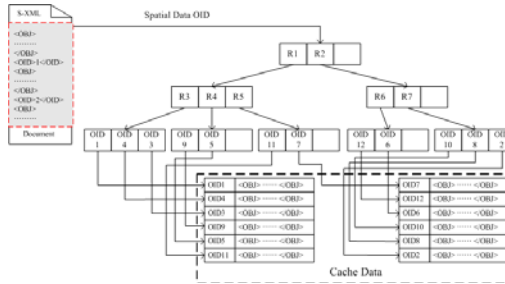


그림 7. R-Tree를 이용한 캐시 데이터 구조

캐시 관리자는 점 질의처리에 대한 결과 공간 데이터의 빠른 검색을 위하여 해시 색인을 갖는다. 해시를 이용한 자료구조는 공간 데이터의 OID가 해시 키 값으로 사용되어 삽입되고, 해시 테이블은 캐시 데이터의 주소 값을 가진다. 이때 R-tree와 해시는 그림 4와 같이 캐시 데이터를 공유한다. 그림 8은 해시를 이용한 캐시 데이터 구조이다.

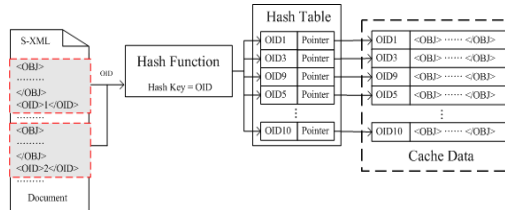


그림 8. 해시를 이용한 캐시 데이터 구조

확장된 S-XML 변환기의 캐시 관리자는 공간 데이터의 계층별로 각각의 색인을 구성한다. 즉 SDBMS의 공간 데이터를 가지는 하나의 테이블이 영역 질의처리를 위한 R-tree와 점 질의처리를 위한 해시 색인을 갖는다. 따라서 S-XML 변환기에서 캐시 데이터를 이용하여 SDBMS에서 처리한 공간 데이터를 변환할 때, 해당되는 색인을 찾기 위한 매핑 정보를 저장하기 위한 자료구조가 필요하다.

SDBMS의 새로운 테이블에서 공간 데이터가 입력되면 캐시 관리자는 해당 테이블의 TID와 공간 데이터를 캐싱한 색인의 포인터를 매핑 테이블에 추가한다. SDBMS의 테이블과 캐시 데이터간의 매핑을 위한 자료구조는 그림 9와 같다.

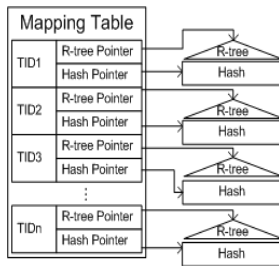


그림 9. SDBMS 테이블과 캐시 데이터 사이의 매핑을 위한 자료구조

3.3 효율적인 S-XML 변환을 위한 캐싱과정

캐시 관리자는 S-XML를 전달받거나 SDBMS의 질의 처리 결과를 변환할 때, 캐시 데이터를 검색하여 변환할 공간 데이터가 캐시에 저장되어 있지 않은 경우 이를 캐싱하여 캐시 관리자의 색인을 통하여 캐시 데이터에 저장한다. 그림 10은 캐시 데이터를 이용한 확장된 S-XML 변환기에서의 질의처리 과정이다.

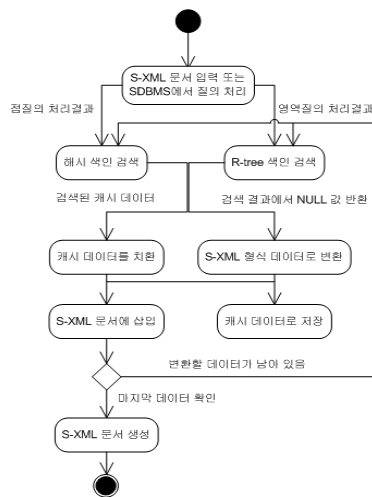


그림 10. S-XML 변환기 질의처리 과정

S-XML 공간 데이터의 캐싱 과정은 점 질의 결과나 영역 질의 결과에 대한 S-XML 데이터에 관계없이 동일한 과정으로 캐싱한다. 하지만 질의 유형에 따른 캐시 데이터의 빠른 검색을 위해 R-tree와 해시 색인을 사용한다.

변환 과정에서 공간 데이터에 대한 변환일 경우 캐시 관리자는 캐시 데이터를 검색하여 S-XML 변환기에 제공하는데, 캐시 데이터에 변환하는 공간 데이터가 존재하지 않으면, 캐시 관리자는 Null 값을 반환하고, S-XML 변환기에서 S-XML 데이터로 변환이 완료되면 해당 데이터는 변환된 형식을 유지하여 캐싱된다. S-XML 변환기에서

변환된 공간 데이터를 MBR을 이용하여 R-tree에 삽입하고, R-tree에서 삽입과정이 완료되면 해시 색인에 OID를 해시 키 값으로 하여 삽입한다. 캐시 데이터는 R-tree와 해시 색인에 각각 저장되는 아니라 캐시 데이터의 포인터를 공유한다. 캐시 관리자에서 S-XML 형태의 공간 데이터 캐싱 알고리즘은 알고리즘 1과 같다.

[알고리즘 1] 공간 데이터 캐싱 알고리즘

```

Input
S-XMLData : 변환된 S-XML 형식의 데이터
OID : 삽입되는 공간 데이터의 OID
MBR : 삽입되는 공간 데이터의 MBR

Begin
01 : if(SearchCacheData(OID) == false)
02 :   InsertInR_tree(MBR, OID, S-XMLData)
03 :   InsertInHash(OID, S-XMLDataPointer)
04 : else
05 :   endToS_XMLTransformation(S-XMLData)
06 : end if

End
    
```

알고리즘 1에서 InsertInR_tree는 캐싱되는 공간 데이터의 MBR과 OID, S-XML로 변환된 데이터를 가지고 R-tree에 삽입하는 함수이고, InsertInHash는 공간 데이터의 OID를 해시 키 값으로 하고, R-tree에서 삽입된 공간 데이터의 포인터를 가지고 해시에 삽입하는 함수이다. 이때 캐시 데이터를 중복하지 않기 위해 해시에서 삽입이 완료되면 입력된 S-XMLDataPointer를 이용하여 R-tree 삽입과정에서 저장된 캐시 데이터를 참조한다.

만일 캐시 데이터를 저장할 메모리 공간이 부족할 경우, 사용횟수가 가장 작은 캐시 데이터를 삭제하여 캐싱에 필요한 공간을 확보한다.

3.4 확장된 S-XML 변환기를 이용한 변환과정

본 장에서는 캐시 데이터를 이용하는 확장된 S-XML 변환기에 대해 설명한다. 변환과정은, 첫 번째로 이진 데이터가 어떤 형태의 질의인지 질의 분석기로부터 정보를 얻어 오는 것이다. 두 번째로 질의 형태에 따라 점 질의에 대한 변환과정과 영역 질의에 대한 변환과정을 진행한다.

영역 질의는 질의에서 요구하는 공간 데이터가 영역을 기준으로 연산될 때, 영역 질의로 판단하고, 캐시 관리자에서 영역 질의의 결과를 위한 R-tree를 사용하여 캐시 데이터를 검색한다. 영역 질의처리 결과에 대한 변환과정은 다음과 같다. SDBMS에서 영역 질의가 처리되어 변환할 이진 데이터가 전송되면, 확장된 S-XML 변환기는 변환하는 결과 데이터가 영역 질의에 의한 것이라는 정보와 질의에서 요구하는 Boundary와 함께 결과 데이터 중에서 공간 데이터의 OID 리스트를 캐시 관리자에 전송한다. 캐시 관리자의 캐싱 오퍼레이터는 R-tree에서 질의 영역을

캐시 데이터를 검색하고, 정제과정을 통하여 캐시 데이터에서 해당되는 공간 데이터와 OID를 임시 버퍼에 저장한다. 확장된 S-XML 변환기는 임시 버퍼에 있는 캐시 데이터를 치환하여 S-XML 문서를 구성하고, 비공간 데이터와 버퍼에 존재하지 않은 공간 데이터를 S-XML 변환기에서 변환하여 S-XML 문서를 생성한다. 영역 질의처리 결과에 대한 S-XML 변환 알고리즘은 알고리즘 2와 같다.

[알고리즘 2] 영역 질의처리 결과에 대한 변환 알고리즘

Input

SBinData : SDBMS에서 처리한 이진 공간 데이터
ASBinData : SDBMS에서 처리한 이진 비공간 데이터
CachedData : 캐시 되어 있는 데이터
OID : 공간 데이터의 OID
Boundary: 질의에서 요구하는 영역 값

Variables

QtyType : 결과 데이터를 요청한 질의의 타입

Begin

```
01 : if(IsRangeQueryType(QtyType) == true)
02 :   Send query information(Boundary) and OID
      list to Cache Manager
03 :   CachedData = R_treeSearch(Boundary)
04 :   TempBuffer = CacheDataFilter(CachedData)
05 :   if(TempBuffer)
06 :     Mapping Spatial Data OID in TempBuffer
07 :     ReplacementToS_XML(TempBuffer)
08 :   end if
09 :   TransformationToS-XML(SBinData)
10 : end if
11 :   TransformationToS-XML(ASBinData)
12 :   Create S-XML document
13 : end if
```

End

알고리즘 2에서 IsRangeQueryType은 입력된 질의가 영역 질의인지 확인하고, R_treeSearch는 R-tree에서 변환에 이용 가능한 캐시 데이터를 찾는 함수이다. CacheDataFilter는 검색된 캐시 데이터에서 질의처리 결과 중 공간 데이터의 OID를 비교하여 질의처리 결과에 해당하지 않는 캐시 데이터를 제거하는 함수이다. CacheDataFilter에서 처리된 캐시 데이터가 임시 버퍼에 저장되면, ReplacementToS_XML 함수는 OID로 매핑하여 임시 버퍼에 저장되어 있는 캐시 데이터를 S-XML 데이터로 치환한다. 마지막으로 TransformationToS-XML은 캐시 데이터에서 검색되지 않은 데이터를 변환하는 기존 S-XML 변환기의 함수이다.

확장된 S-XML 변환기에서는 점 질의를 질의에서 요구하는 공간 데이터의 형식이 영역을 기반으로 한 질의 이외의 모든 질의를 점 질의로 판단한다.

점 질의에 대한 결과를 S-XML 데이터로 변환하는 과

정은 캐시 관리자의 해시를 사용하여 캐시 데이터를 검색하고, 검색된 캐시 데이터를 확장된 S-XML 변환기에서 치환하여 사용함으로써 공간 데이터에 대한 변환 횟수를 감소할 수 있다. 점 질의의 결과를 변환하는 자세한 과정은 다음과 같다.

첫 번째로 변환 될 이진 데이터가 입력되면 S-XML 변환기는 질의 정보와 공간 데이터의 OID 리스트를 캐시 관리자에 전송한다. 캐시 관리자의 캐싱 오퍼레이터는 질의형태가 점 질의 일 경우 해시를 사용하여 전송 받은 OID 리스트에 해당되는 공간 데이터가 캐시 된 데이터에 존재하는지 검색한다. 검색을 통하여 캐시 된 데이터 중에서 해당되는 데이터는 S-XML 변환기로 전송되고, 캐시 된 공간 데이터는 S-XML 형식의 스트링 데이터이므로 S-XML 변환기는 이를 치환 하여 사용한다. 캐시 된 데이터에서 검색 되지 않은 공간 데이터와 비공간 데이터는 S-XML 변환기에서 처리하여 S-XML 문서를 생성한다. 점 질의처리 결과에 대한 S-XML 변환 알고리즘은 알고리즘 3과 같다.

[알고리즘 3] 점 질의처리 결과에 대한 변환 알고리즘

Input

SBinData : SDBMS에서 처리한 이진 공간 데이터
ASBinData : SDBMS에서 처리한 이진 비공간 데이터
CachedData : 캐시 되어 있는 데이터
OID : 공간 데이터의 OID

Variables

QtyType : 결과 데이터를 요청한 질의의 타입

Begin

```
01 : if(IsPointQueryType(QtyType) == true)
02 :   Sendquery information and OID list to Cache
      Manager
03 :   CachedData = HashSearch(OID)
04 :   if(Search CacheData)
05 :     Send CachedData to S-XML Transformation
      Component
06 :     ReplacementToS_XML(CachedData)
07 :   end if
08 :   TransformationToS-XML(SBinData)
09 :   TransformationToS-XML(ASBinData)
10 :   Create S-XML document
11 : end if
```

End

알고리즘 3에서 IsPointQueryType은 변환 중인 공간 데이터가 점 질의처리에 의한 것인지 확인하는 함수이고, HashSearch는 캐시 관리자의 해시 색인에서 변환 중인 공간 데이터의 OID를 이용하여 캐시 데이터를 검색하는 함수이다. ReplacementToS_XML 함수는 확장된 S-XML 변환기에서 전송 받은 캐시 데이터를 S-XML 데이터로 치환한다. TransformationToS-XML 함수는 알고리즘 2의 영역 질의처리 결과에 대한 S-XML 변환 알고

리즘에서와 같다.

4. 성능평가

본 장에서는 S-XML 관리 시스템에서 기존의 S-XML 변환기와 확장된 S-XML 변환기의 변환성을 비교 평가한다. 기존의 S-XML 변환기에서는 입력 질의가 SDBMS에서 처리되면, 질의처리 결과를 전송 받아 S-XML 문서로 변환한다. 반면에 확장된 S-XML 변환기에서는 비공간 데이터에 비해 변환 비용이 큰 공간 데이터의 변환 결과를 캐싱하여 변환과정에 이용한다.

실험환경으로는 2.4GHz(L1 64kb, L2 8MB) CPU, 4GB 메모리를 사용하였으며, 운영체제는 MS windows XP professional, 개발 언어는 C++를 사용하였다.

평가 항목은 변환하는 데이터의 크기에 따른 변환 시간과, 변환에 이용되는 데이터들의 캐시 적재율에 따른 변환 시간을 측정하였다.

첫 번째, 변환 데이터의 크기에 따른 성능평가로 동일한 조건에서 변환되는 데이터의 크기를 증가 시키면서 변환 시간을 실험하였다. 데이터의 크기는 변환되는 공간 데이터의 라인 스트링을 구성하는 포인트의 개수를 의미한다. 즉, 라인 스트링을 구성하는 포인트의 개수가 많을 수록 크기가 큰 데이터이다. 이 장에서 변환되는 데이터의 크기를 기준으로 하여 변환 성능을 평가하는 것은, 공간 데이터가 여러 종류의 타입과 복잡한 구조를 이루어져 있지만 실질적으로는 포인트들을 가지는 라인 스트링으로 표현되기 때문이다.

실험에 사용된 데이터는 데이터 개수는 50000개이며, 데이터 크기는 라인 스트링을 이루는 포인트 수를 1부터 200개까지 10개씩 증가하였다. 데이터 크기에 따른 변환 성능 실험 결과는 그림 11과 같다.

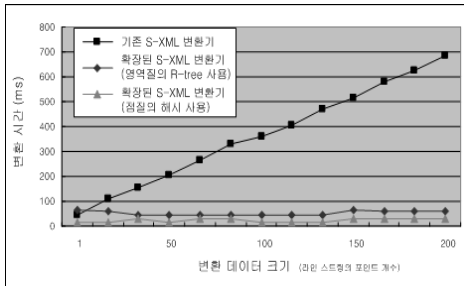


그림 11. 변환 데이터 크기에 따른 성능평가 결과

데이터 크기에 따른 변환 성능평가에서 기존 변환과정은 데이터의 크기가 증가함에 따라 변환 연산비용이 증가하였고, 이에 변환 시간이 선형적으로 증가하였다. 확장된 S-XML 변환기의 캐시 데이터를 이용한 변환에서는 이미 S-XML 형태로 변환되어 있는 캐시 데이터를 치환함으로써 변환 연산과정이 필요하지 않아, 변환하는 데이터 크기에 관계 없이 거의 일정한 변환 시간을 보이고 있다.

다음으로 캐시 적재율에 따른 변환 성능 평가에서는 동일한 조건에서 변환과정에 이용할 수 있는 공간 데이터의 캐시 적재율을 증가 시키면서 변환 시간 측정하였다. 실험에 사용된 데이터는 100개의 포인트로 구성된 라인 스트링 10000개로 동일하며, 캐시 적재율은 0%에서 100%까지 증가 시키면서 실험하였으며, 캐시 적재율 0%는 기존 S-XML 변환과 같다. 캐시 적재율에 따른 변환 성능 실험 결과는 그림 12와 같다.

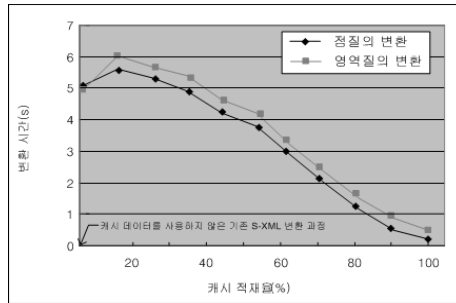


그림 12. 캐시 적재율에 따른 성능평가 결과

캐시 적재율에 따른 성능평가에서는 적재율이 30% 미만일 경우 기존 S-XML 변환기와 같은 변환과정이 이루어지는 0%의 경우보다 높은 변환 시간이 측정되었는데 이는 본 성능평가에서 캐시 적재율에 따라 메모리 크기를 제한하여, 필요한 캐시 데이터가 검색되지 않는 경우 삭제, 변환, 캐싱 과정에서 발생하는 비용이 캐시 데이터를 사용하는 이익 비용보다 크기 때문이다. R-tree를 이용한 영역 질의처리에 대한 변환 시간은 캐시 적재율이 35%를 넘으면서 기존 변환기보다 우수한 성능을 보였고, 70%를 상회하면서 기존 변환기의 변환 시간은 50%이상 줄었다. 해시 색인을 이용하는 점 질의처리 결과의 변환의 경우 캐시 적재율이 25%를 넘으면서 기존 변환기의 변환시간 보다 빠른 변환을 하였고, 65% 를 상회하면서 기존 변환기보다 50%이상의 성능을 보였다.

5. 결론 및 향후연구

u-GIS 환경에서 웹 기반의 공간정보 응용서비스들은 공간 데이터의 유통을 위한 표준 형식으로 S-XML, GML과 같이 공간 데이터 스키마를 지원하는 확장된 XML을 사용한다. 본 논문에서는 S-XML의 인터넷 유통 과정 중 SDBMS와의 효율적인 변환 기법을 제안하였다. 이는 S-XML 데이터에서 공간 데이터에 해당하는 부분만을 캐싱하고, 상호 변환 과정에서 캐시 데이터를 이용하여 공간 데이터에 대한 부하를 줄임으로써 S-XML 관리 시스템의 변환 성능을 향상시켰다. 제안 기법은 S-XML 관리 시스템에서 공간 데이터에 대한 질의를 점 질의와 영역 질의로 나누고 캐시 관리자에서 점 질의에 대한 결과 S-XML에서의 공간 데이터는 해시 색인을 사용하여 캐시

데이터를 저장, 검색하였고, 영역 질의에 대한 결과 데이터는 R-tree를 사용하였다.

성능 평가에서는 공간 데이터의 크기와 변환과정에 사용되는 공간 데이터의 캐시 적재율을 기준으로 하여 측정하였다. 성능 평가 결과에서 공간 데이터의 크기가 늘어남에 따라 기존 S-XML 변환기에서는 변환 시간이 늘어났다. 이 실험의 결과에서 알 수 있듯이 복잡한 형태의 대용량 공간 데이터의 변환과정에서 제안 기법이 효과적임을 증명하였다. 또한 캐시 적재율에 따른 성능평가에서는 변환에서 사용되는 공간데이터가 30%이상 캐시에 적재되어 있는 경우 기존 S-XML 변환기 보다 우수한 변환 성능을 보여 실제 운영되는 시스템에서도 제안된 S-XML 변환기의 가능성을 보였다. 향후 연구로는 공간 정보의 특성을 활용하여 캐시 데이터의 재사용성을 높이기 위한 연구가 필요하다.

참고 문헌

- [1] 이충호, 안경환, 이문수, 김주완, "u-GIS 공간정보 기술 동향," 전자통신동향분석 제22권 3호, 2007.
- [2] T. Bray, J. Paoli, S. McQueen, and C.M., "Extensible Markup Language (XML) 1.0 (2nd Edition)," W3C Recommendation, 2000.
- [3] W3C, XML Specification ("XMLspec"), 1998.
- [4] OGC, Conformance Test Guidelines for Open GIS Simple Features Specification for SQL, Revision 1.0, 1998.
- [5] OGC, "Geography Markup Language (GML) Implementation Specification 3.0," 2003.
- [6] G-XML: <http://gisclh.dpc.or.jp/gxml/>
- [7] 안영수, 박순영, 정원일, 배해영, "공간 데이터베이스와 GML 문서간의 지리 정보 변환 방법," 한국정보처리학회 추계학술대회 논문집, 2002.
- [8] 이동욱, 장용일, 박순영, 오영환, 배해영, "S-XML 미들웨어에서 캐싱을 이용한 효율적인 공간 데이터 변환 기법," 한국정보처리학회 춘계학술발표대회 논문집, 2005.
- [9] D. Lee, M. Mani, and W. Chu, "Effective Schema Conversions between XML and Relational Models," In European Conf. on Artificial Intelligence (ECAI), Knowledge Transformation Workshop (ECAI-OT), Lyon, France, July 2002.
- [10] D. Calvanese, G. Giacomo, M. Lenzerini, and M. Vardi, "Answering Regular Path Queries Using Views," Proc. of the IEEE Int'l Conf. pp.389-398, 2000.
- [11] A. Levy, A. Medelzon, Y. Sagiv, and D. Srivastava, "Answering Query Using Views," Proc. of the ACM Int'l Symp, pp.95-104, 1995.
- [12] S. Abiteboul, L. Segouin, and V. Vianu, "Representing and Querying XML with Incomplete Information," Proc. of the ACM Int'l Symp. on PODS, pp.150-161, 2001.
- [13] V. Hristidis and M. Petropoulos, "Semantic Caching of XML Database," Proc of the Int'l Workshop on the Web and Databases, 2002.
- [14] 문찬호, 박정기, 강현철, "실체뷰 캐시 기법을 이용한 XML 질의 처리 시스템의 구현," 한국정보처리학회 논문지D, 2004.
- [15] 안 영수, 박순영, 정원일, 배해영, "GML 문서의 통합 지리 정보 검색을 위한 XQuery의 확장," 한국정보과학회 춘계학술대회 논문집, 2003.



이 동 옥

2003년 상지대학교 전자계산공학과 (이학사)
2005년 인하대학교 컴퓨터 정보공학과 (공학석사)
2005년~현재 인하대학교 정보공학과 (박사과정)

관심분야는 유비쿼터스 환경을 위한 공간 DBMS 및 DSMS



백 성 하

2005년 인하대학교 수학통계학부 (이학사)
2007년 인하대학교 컴퓨터 정보공학과 (공학석사)
2007년~현재 인하대학교 정보공학과 (박사과정)

관심분야는 데이터 스트림 관리 시스템, 데이터베이스 클러스터



김 경 배

1992년 인하대학교 전자계산공학과 (공학사)
1994년 인하대학교 전자계산공학과 (공학석사)
2000년 인하대학교 전자계산공학과 (공학박사)

2004년~현재 서원대학교 컴퓨터교육학과 조교수

관심분야는 이동실시간 데이터베이스, 스토리지 시스템, GIS

배 해 영



1974년 인하대학교 응용물리학과 (공학사)
1978년 연세대학교 전자계산학과 (공학석사)
1989년 숭실대학교 전자계산학과 (공학박사)

1982년~현재 인하대학교 정보공학부 교수

1999년~현재 인하대학교 지능형GIS연구센터 센터장

관심분야는 분산 데이터베이스, 공간 데이터베이스, 지리정보 시스템, 멀티미디어 데이터베이스 등