

Linux 기반 Shareware DBMS를 이용한 공간 데이터베이스 시스템의 설계 및 구현[†]

윤재관⁰ 장영승 한기준
건국대학교 컴퓨터공학과

Design and Implementation of a Spatial Database System Using Shareware DBMS based on Linux

Jae-Kwan Yun, Yan-Sheng Zhang, and Ki-Joon Han
Department of Computer Science & Engineering, Kon-Kuk University

요 약

최근 들어 지리 정보 시스템(GIS)의 사용이 점차 다양화되고, 점차적으로 다양한 계층의 사용자가 접근을 하게 되면서 뛰어난 성능의 지리 정보 시스템이 점차적으로 요구되고 있다. 그러나, 이러한 고성능의 지리 정보 시스템의 사용은 고가의 운영체제와 지리 정보 시스템을 요구하고 있기 때문에 일반적인 형태의 서비스에는 쉽게 사용되기 어렵다. 본 논문에서는 저가의 시스템 비용으로 지리 정보 시스템을 구축하기 위해 점차적으로 많이 사용되고 있는 저가의 안정적인 시스템인 Linux 환경에서 경량급이면서 성능이 뛰어난 Shareware DBMS를 사용하여 공간 데이터베이스 시스템을 설계 및 구현한다. 본 논문에서 제안하는 공간 데이터베이스 시스템은 사용자 응용 프로그램이나 WEB 인터페이스에서 본 시스템에 접근하기 위한 인터페이스 관리자, 요청된 질의를 처리하기 위한 질의 관리자, 공간 데이터의 획득 및 공유를 위한 수입/수출 관리자, 공간 데이터를 처리하기 위한 공간 연산 관리자, 공간 데이터에 대한 Collection을 처리하기 위한 Collection 관리자, 그리고 공간 데이터의 빠른 검색을 위한 인덱스 관리자로 구성된다.

1. 서론

최근 컴퓨터의 대용량화 및 고성능화에 따라 지리 정보 시스템에 근거한 다양한 형태의 응용 시스템이 연구 및 개발되고 있다 [3, 6, 7]. 이전까지의 이러한 연구 및 개발은 주로 지리 정보 시스템에 대한 전문가를 기준으로 많이 진행되다가 인터넷의 점진적인 발전으로 인하여 WWW을 중심으로 한 일반 사용자 측면의 기술로 발전되어 가고 있다. 즉, 이제까지의 지리 정보 시스템은 도로망, 상하수도관 등과 같은 특수 형태의 업무에 대한 데이터 처리에 주로 사용되었지만 점차적으로 일반 사용자를 위해 교통 현황 정보 제공, 지도 검색 등과 같은 다양한 분야에서 사용되어 가고 있다 [5, 8].

이렇게 지리 정보 시스템에 대한 수요가 급격히 증가하는 가운데 대부분의 지리 정보 서비스 사업자들은 값비싼 유닉스 시스템과 고가의 지리 정보 시스템을 이용하여 서비스를 하고 있다. 그러나, 점차적으로 저가의 운영체제인 Linux의 사용이 확대되어 가고 있고, Linux 기반의 응용 프로그램이 증가하고 있기 때문에 Linux 기반의 지리 정보 시스템의 개발이 필요한 시점이다. 이러한 Linux 기반의 지리 정보 시스템을 사용하게 되면 지도 데이터의 서비스를 위한 시스템 구축에 사용되는 비용을 절감시킬 수 있고, 또한 지리 정보 시스템의 사용도 보다 더 확산시킬 수 있다.

이를 위해서 본 논문에서는 안정성 및 효율성을 기반으로 하여 빠르게 확산되는 운영체제인 Linux 환경에 적합하고, 클라이언트/서버 구조 및 인터넷 환경을 지원하는 Shareware DBMS를 이용하여 공간 데이터베이스 시스템을 설계 및 구현한다.

본 논문의 구성은 다음과 같다. 제 1장 서론에 이어 제 2장에서는 현재 Linux에서 사용되고 있는 Shareware DBMS와 공간 데이터 교환을 위한 표준인 SDTS에 대하여 설명한다. 그리고, 3장에서는 전체적인 시스템의 구조 및 공간 데이터의 구조, 공간 연산자, 그리고 Collection에 대하여 설명한다. 제 4장에서는 인터페이스 관리자, 질

의 관리자, 인덱스 관리자, 공간 연산 관리자, Collection 관리자, 그리고 수입/수출 관리자의 인터페이스에 대하여 살펴본다. 마지막으로 제 5장에서는 결론 및 향후 연구 과제에 대하여 언급한다.

2. 관련 연구

본 장에서는 현재 Linux 기반의 Shareware DBMS의 특성에 대하여 살펴보고, 본 논문에서 공간 데이터를 수입/수출하기 위해 사용한 SDTS에 대하여 살펴본다.

2.1 Linux 기반 Shareware DBMS

현재 Linux 기반의 Shareware DBMS는 PostgreSQL, MySQL, MySQL, Oracle for Linux, Informix for Linux 등이 있다. PostgreSQL은 강력한 객체 관계형 데이터 모델과 풍부한 데이터 타입을 지원하고, 고수준의 확장성을 가지고 있으며, 복합 객체 (complex object), 규칙(rule) 등의 기능과 질의 최적화, 동시성 제어, 트랜잭션 처리, 다중 사용자 기능을 제공하고 있다. MySQL은 경량급의 관계형 데이터베이스 시스템이며 클라이언트 응용 프로그램을 위해 C, Java, Perl5, ESL, Tcl/Tk, Python, C&C++ 등의 인터페이스 라이브러리를 제공하고 있다. MySQL도 관계형 데이터베이스 시스템이며, 커널 쓰레드를 이용한 멀티 쓰레드를 지원하므로 CPU가 여러 개일 경우에 효율성이 높다. 그리고, 다양한 플랫폼을 제공하고 있으며, 빠르고 안정적이다. Informix for Linux는 효율적인 데이터 저장 방법을 제공함으로써 빠른 데이터 접근이 가능하고 멀티 프로세스를 지원할 수 있다. 그리고, 분산형 클라이언트/서버 기능을 제공하고 있으며, 다양한 형태의 데이터 타입을 제공하고 있다. Oracle for Linux는 다중 쓰레드 방식을 지원하고 있으며 행단위 잠금 방식을 사용한다. 그리고, 데이터베이스 화일의 크기를 자유롭게 바꿀 수 있으며, 병렬 질의 처리 기능을 지원하고, 시스템이 동작하고 있는 도중에도 백업과 복구가 가능하다.

Linux 기반의 Shareware DBMS는 대체로 경량급의 형태를 취하고 있지만 다양한 플랫폼을 제공하고, 데이터의 빠른 접근을 허용하

[†] 본 연구는 한국통신정보기술(주)의 위탁연구과제에서 지원받았음.

고, 또한 안정적이고 가격이 싼 장점이 있다 [4]. 그러므로, 이러한 Shareware DBMS를 지리 정보 시스템에 적용하여 사용하게 되면 낮은 시스템 구축 비용으로 충분한 기대 효과를 볼 수 있다. 본 논문에서는 이러한 Linux 기반의 Shareware DBMS를 이용하여 안정적이고 효율적인 공간 데이터베이스 시스템을 설계 및 구현한다.

2.2 SDTS

SDTS는 다양한 사용자들이 지리 데이터 구축의 중복성을 피하고, 효율적으로 지리 데이터를 공유할 목적으로 등장한 공통 데이터 교환 포맷이다 [2]. SDTS는 현재 6개의 Part로 나뉘어져 있으며 그 중 Part 1, 2, 3은 일반적인 정의 및 규정에 관한 내용이고, Part 4, 5, 6은 프로파일(Profile)에 관한 내용이다.

Part 1은 공통 데이터 교환의 논리적, 개념적 규약으로 공간 데이터의 개념적 모델, 데이터 질 부분, 변환과 관련된 레코드와 필드 수준의 내용을 포함한다. Part 2는 실제계 공간 현상들, 속성들, 속성값들의 정의 모델을 제공하고, 이들의 표준화된 용어 리스트를 포함한다. Part 3은 물리적 데이터 변환 방식 표준인 ISO/ANSI 8211(FIPS 123)을 사용하여 Part 1의 논리적 규약을 구체적인 물리적 수준으로 전환 가능하도록 규약한다. Part 4, 5, 6은 Part 1, 2, 3의 일반 규약을 토대로 구체적인 데이터 전환 규약을 가진 프로파일을 정의한다. 현재 Part 4는 벡터 데이터의 규정 및 변환을 위한 TVP(Topological Vector Profile), Part 5는 래스터 데이터를 위한 RP(Raster Profile), Part 6은 점 데이터를 위한 PP(Point Profile)로 각각의 프로파일을 기술한다. 본 논문에서는 공간 데이터베이스 시스템에서 사용되는 공간 데이터의 수입/수출을 위하여 SDTS를 사용한다.

3. 시스템의 설계

본 장에서는 Linux 기반의 Shareware DBMS를 이용한 공간 데이터베이스 시스템의 전체적인 구성 및 본 시스템에서 제공하는 공간 데이터 타입, 공간 연산자, 그리고 Collection에 대하여 설명한다.

3.1 전체적인 시스템의 구조

본 논문에서 제시하는 공간 데이터베이스 시스템은 인터페이스 관리자, 질의 관리자, 수입/수출 관리자, 공간 연산 관리자, Collection 관리자, 그리고 인덱스 관리자의 6개 관리자로 나뉘어진다. 그림 1은 본 논문에서 제시하는 공간 데이터베이스 시스템의 전체 구조를 보여준다.

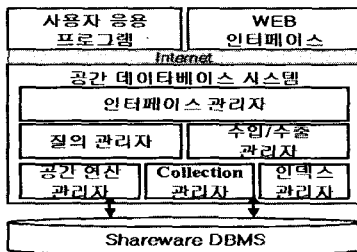


그림 1. 시스템의 전체 구조

지리 정보 시스템은 사용자 응용 프로그램이나 WEB 인터페이스를 통해 인터페이스 관리자의 API를 이용하여 공간 데이터베이스 시스템에게 데이터를 요청하게 된다. 인터페이스 관리자는 이러한 데이터의 요청을 질의 관리자에게 전달하고, 질의 관리자가 반환한 질의 결과를 전달한다. 질의 관리자는 질의의 형태에 따라서 Collection에 대한 처리를 수행하는 Collection 관리자, 인덱스에 대한 작업을 수행하는 인덱스 관리자, 공간 연산을 처리하는 공간 연산 관리자를 이용하여 질의를 처리하게 된다. 그리고, 수입/수출 관리자는 공간 데이터를 공유 및 확보하기 위하여 사용된다.

3.2 공간 데이터의 타입

본 논문에서 개발할 Linux 기반 Shareware DBMS를 이용한 공간 데이터베이스 시스템에서는 표 1과 같은 3가지의 공간 데이터 타입을 제공한다. 점 객체는 일반적으로 point(x,y)로써 표현되고, 선 객체는 2개의 점을 나타내는 값으로써 표현된다. 그리고, 다각형 객체는 3개 이상의 점 객체로써 표현되며 점의 개수와 사이즈를 이용하

여 몇 개의 점 객체로 이루어졌는지 알 수 있다. 이러한 공간 데이터는 공간 데이터베이스 시스템에 저장될 때 사용되는 Shareware DBMS에 따라 공간 데이터 타입의 형태로 저장되는 경우와 문자열의 형태로 저장되는 경우로 구분된다. 문자열 형태로 저장된 공간 데이터의 형태는 StrToPoint, StrToLine, StrToPolygon, PointToStr, LineToStr, PolygonToStr 등의 데이터 변환 함수에 의하여 공간 데이터 타입의 형태로 변환된다.

표 1. 공간 데이터의 타입

공간 데이터 타입	이름	타입	설명
TPoint	x	double	x 좌표
	y	double	y 좌표
TLine	p[0]	TPoint	시작점
	p[1]	TPoint	끝점
TPolygon	size	int	길이
	npts	int	점의 개수
	pts	TPoint	(x[0], y[0], x[1], y[1], .. x[npts-1], y[npts-1])

3.3 공간 연산자

질의 관리자에서 공간 데이터에 대한 연산을 수행하기 위해서는 공간 데이터에 대한 공간 연산자를 제공하여야 한다. 그림 2는 공간 데이터베이스 시스템에서 지원하는 공간 데이터의 처리를 위한 공간 연산자들을 보여준다. CONTAIN과 CONTAINED는 2개의 공간 객체가 서로 겹쳐 있는지에 관련된 공간 연산자이고, COVER과 COVERED는 하나의 공간 객체가 다른 공간 객체를 포함하거나 포함되어 있는지에 관련된 공간 연산자이다. 또한, CROSSOVER는 서로 다른 차원의 공간 객체 중 높은 차원의 공간 객체가 다른 낮은 차원의 공간 객체 위를 지나가는 것에 관한 공간 연산자이다. DISJOINT는 두 공간 객체 사이에 포함 관계가 없는 경우이고, OVERLAP은 두 개의 공간 객체가 겹쳐있는 경우에 관한 공간 연산자이다. TOUCH는 공간 객체 사이에 내부적으로 겹치는 곳이 없이 외부적으로 겹쳐진 상태에 관한 공간 연산자이다.

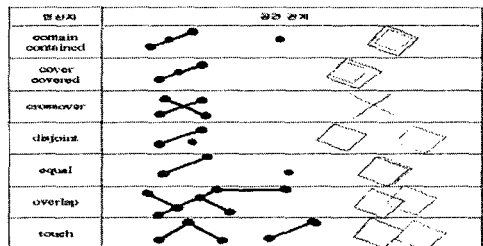


그림 2. 공간 연산자

3.4 Collection

같은 범주에 속하는 공간 데이터를 처리하기 위해서는 Collection을 지원하여야 한다. 예를 들어, "광진구" 지역에 해당하는 모든 "학교" 객체에 대한 검색 연산을 수행하고자 할 때, 합수로써 해당하는 객체를 찾아내기 위해서는 시간이 많이 소요된다. "학교"는 "초등학교", "중학교", "고등학교", "대학교" 등의 여러 개의 클래스에 해당하는 객체이다. 그러므로, 사용자가 해당하는 객체들을 미리 검색하여 별도의 공간에 저장하는 관리 기법이 필요하다. 이러한 Collection에 대한 연산은 관계형 데이터베이스 시스템에서의 기본키, 객체 지향 데이터베이스 시스템과 객체 관계형 데이터베이스 시스템의 객체 식별자를 이용하여 처리된다.

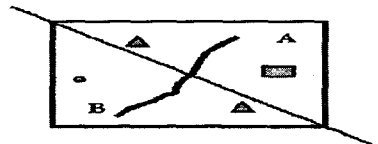


그림 3. Collection의 예

공간 데이터베이스 시스템에서는 Collection에 대해서 List와 Set을 지원한다. List는 중복과 순서를 가지고 있는 객체들의 집합이고, Set은 중복성과 순서를 가지지 않는 객체들의 집합이다. Collection의 처리는 그림 3에서와 같이 전체를 A와 B로 구역을 나누었을 때 삼각형으로 표현된 객체는 같은 테이블에서 검색된 공간 데이터이지만 구역에 따라서 다른 Collection에 저장된다. 또한, A와 B에 걸쳐서 표현되는 공간 데이터는 내부의 Collection 처리 함수에 의해 분리되어 저장된다.

4. 시스템의 구현

본 장에서는 Linux 기반 Shareware DBMS를 이용한 공간 데이터베이스 시스템을 구성하는 인터페이스 관리자, 질의 관리자, 인덱스 관리자, 공간 연산 관리자, Collection 관리자, 그리고 수입/수출 관리자의 구현에 대해 설명한다.

4.1 인터페이스 관리자

인터페이스 관리자는 사용자 응용 프로그램이나 WEB 인터페이스에서 공간 데이터베이스 시스템에 접근하기 위한 API를 제공한다. 이러한 API는 Shareware DBMS에서 제공하는 데이터를 처리하기 위한 함수들을 이용하여 구성되며 TPostgres, TMySQL 등과 같은 클라이언트를 추가함으로써 다른 데이터베이스에 접근할 수 있게 된다. 인터페이스 관리자에서는 데이터베이스 객체를 생성하고 소멸하기 위한 생성자와 소멸자가 있으며, 연결을 위한 ConnectToDB, DisconnectFromDB 등의 함수가 있다.

4.2 질의 관리자

질의 관리자는 인터페이스 관리자를 통하여 요청된 질의를 처리한다. 질의 관리자에서 처리하는 질의는 공간 및 비공간 데이터에 대한 삽입, 삭제, 갱신 및 검색 등이 있다. 질의 관리자에서 제공하는 함수는 트랜잭션을 처리하기 위한 BeginTransaction, CommitTransaction, AbortTransaction, 질의를 수행하기 위한 ExecSQL, 커서를 설정하기 위한 DECCUR, 설정된 커서를 이용하여 데이터를 로드하기 위한 FETCHALL, 커서를 종료하기 위한 CLOSECUR 등의 함수가 있다. 그리고, 검색된 결과에 대한 정보를 얻기 위한 GetFieldNum, GetRecordNum, GetFieldName, GetValueByName 등의 함수와 레코드를 이동하기 위한 First, Last, Next, Prev 등의 함수가 있다.

4.3 인덱스 관리자

질을 처리하기 위해서는 논리적으로나 물리적으로 엄청난 양의 데이터를 메모리로 적재하여 검색하는 과정이 필요하게 된다. 이렇게 대용량의 데이터를 효율적으로 처리하지 않으면 시스템의 성능이 상당히 저하된다. 본 논문에서의 인덱스 관리자는 효율적으로 공간 및 비공간 데이터를 검색하기 위해서 사용되는데 비공간 데이터의 빠른 검색을 위해서 B+tree를 사용하고, 공간 데이터의 검색을 위해서는 R-Tree, R*-tree를 사용한다 [1]. 인덱스 관리자에서 제공하는 함수는 Create(A)SIndex, Destroy(A)SIndex, Insert(A)SIndex, Remove(A)SIndex, Update(A)SIndex 등의 함수가 있다.

4.4 공간 연산 관리자

공간 연산 관리자는 조건에 맞는 공간 데이터를 검색하기 위해서 필요한 공간 연산자들을 제공한다. 공간 연산 관리자의 각 함수는 질의 관리자에서 의하여 호출되며 정확한 공간 데이터를 효율적으로 검색하기 위해서 사용된다. 공간 연산에 사용되는 함수들은 2차원적 공간 연산에 관한 함수이며, 공간 데이터의 타입에 따라 구분된다. 본 논문에서는 PointContain(ed), PointEqual, PointTouch 등의 점에 관한 공간 연산, LineEqual, LineOverlap, LineTouch, LineContain(ed), LineCover(ed) 등의 선에 관한 공간 함수, PolygonContain, PolygonCover(ed), PolygonCrossover 등의 다각형에 관한 공간 함수를 제공한다.

4.5 Collection 관리자

공간 데이터베이스 시스템에서 많이 사용되는 기능 중의 하나는 Collection이다. 사용자에게 의해 검색되는 공간 데이터는 실제 필요한 데이터 외에 관계없는 데이터가 많이 포함된다. 예를 들어, 사용자가

원하는 위치를 알려주기 위해서는 그 주변의 공간 데이터까지 같이 검색되는 경우가 많다. 이러한 경우에 질의에 사용되는 데이터가 많기 때문에 공간 데이터베이스 시스템의 성능을 상당히 저하시키게 된다. Collection 관리자는 질의에 사용되는 공간 데이터를 위하여 TList와 TSet 클래스를 제공한다. 각 클래스의 멤버 함수로는 Create, Destroy, Add, Delete, IndexOf, Sort 등의 함수가 있다.

4.6 수입/수출 관리자

본 논문에서 개발한 공간 데이터베이스 시스템에서는 실제적으로 사용될 수 있는 공간 데이터를 수집하고 다른 시스템과 데이터를 공유하기 위하여 수입/수출 관리자를 제공하고 있다. 수입/수출 관리자에서는 국가공통 교환포맷인 SDTS 데이터에 대한 수입/수출 기능을 제공한다. 수입 관리자에서는 SDTS 데이터를 수입하기 위하여 중간 파일을 생성하고, 텍스트 형태로 생성된 중간 파일을 Linux의 화일 시스템을 이용하여 값을 읽어 들인 후 공간 데이터베이스 시스템에 저장한다. 또한, 수출 관리자에서는 공간 데이터를 중간 화일 형태로 출력하고, 생성된 중간 화일은 SDTS의 모듈에 맞게 SDTS 데이터로 변환된다. 수입/수출 관리자에서 제공하는 함수는 ExportToFile, ImportFromFile 등의 함수가 있다.

5. 결론 및 향후 연구 과제

지리 정보 시스템의 사용이 점차적으로 일반화 되면서 보다 좋은 지리 서비스를 값싼 가격으로 제공하기 위한 방법이 필요하게 되었다. 이를 위해서는 우선 안정적인 운영체제를 기반으로 하여 효율적인 지리 정보 시스템을 개발하는 것이 중요하다. 이에 본 논문에서는 Linux 운영체제를 기반으로 하여 Shareware DBMS를 이용한 공간 데이터베이스 시스템을 설계 및 구현하였다.

본 논문에서 개발한 공간 데이터베이스 시스템은 다양한 형태의 Shareware DBMS를 이용하여 공간 및 비공간 데이터를 저장, 검색, 삭제, 갱신할 수 있으며, Collection과 공간 연산자 등과 같은 공간 데이터 처리에 필요한 기능들을 제공하고 있다. 또한, 사용자 응용 프로그램이나 WEB 인터페이스로부터 공간 데이터베이스 시스템에 접근하기 위한 인터페이스 관리자, 공간 및 비공간 데이터에 대한 질의를 수행할 수 있는 질의 관리자, 빠른 데이터의 검색을 위한 인덱스 관리자, 그리고 공간 데이터에 대한 수입/수출 기능을 제공하는 수입/수출 관리자 등을 제공한다.

향후 연구 과제로는 공간 데이터에 대한 효율적인 분배 및 검색을 위해 분산 환경을 지원하는 공간 데이터베이스 시스템을 개발하는 것이다.

참고문헌

- [1] Beckmann, N., Kriegel, H.P., Schneider, R., and Seeger, B., "The R*-tree: An Efficient and Robust Access method for Points and Rectangles," Proc. of the ACM SIGMOD Int. Conf., 1990, pp. 322-331.
- [2] National Institute of Standards and Technology, *The Spatial Data Transfer Standard, Federal Information Processing Standard Publication 173*, U.S. Department of Commerce, 1992.
- [3] Oh, B.W., and Han, K.J., "GOOD : A Geographical Data Manager Using Spatial Indices," Proc. of the Int. Workshop on Issues and Applications of Database Technology, Germany, 1998, pp.425-431.
- [4] PostgreSQL Inc, <http://www.postgresql.org>.
- [5] 강홍근, 민준기, 이성진, 정진환, "OMEGA 공간 객체 관리 시스템의 객체 관리기의 설계 및 구현," 데이터베이스연구회, 데이터베이스 학술대회 논문집, 14권1호, 1998, pp. 197-202.
- [6] 윤재관, 오병우, 한기준, "공간 데이터 마이닝을 위한 객체 관리 시스템," 한국정보과학회 학술발표 논문집, 25권1호, 1998, pp.36-38.
- [7] 윤재관, 오병우, 한기준, "공간 데이터 마이닝을 위한 개방형 객체 관리 시스템의 설계 및 구현," 한국개방형GIS연구회 논문지 1권1호, 1999, pp.5-18.
- [8] 이강준, 강태영, 한기준, "GORD : 지리 객체-관계형 데이터베이스 관리 시스템," 한국정보과학회 학술발표 논문집, 24권1호, 1997, pp.167-170.