

비공유 공간 데이터베이스 클러스터에서 한정된 자원에 적합한 형태로의 공간데이터 처리 기법

유병섭^o 장용일 이충호 배해영
인하대학교 전자계산공학과
{subi^o, himalia, chlee}@dblab.inha.ac.kr, hybae@inha.ac.kr

The Mechanism of Spatial Data Processing for Restrictive Resource-Adaptive Form in Shared-nothing Spatial Database Cluster

Byeongseob You^o Yongil Jang Chungho Lee Haeyoung Bae
Dept. of Computer Science and Engineering, INHA University

요 약*

최근 모바일 장치의 발전과 보급으로 차량 네비게이션 시스템(navigation system)과 같이 모바일 장치를 이용한 지도정보 서비스가 발전하였다. 모바일 장치는 자원이 한정되어 있고, 네트워크의 대역폭이 적기 때문에 크기가 크고 복잡한 벡터 데이터를 효율적으로 전송할 수 있는 방법이 매우 중요하다. 본 논문에서는 비공유 데이터베이스 클러스터에서 한정된 자원에 적합한 형태로 공간데이터를 처리하는 기법을 제안한다. 이는 스케일과 사용자의 요청에 따라 보여질 필요가 없는 피쳐들을 제거하고 스케일에 따라 보여지는 형태가 바뀌는 피쳐들에 대해 공간타입을 바꾸는 필터링을 통해 최적화된 지도 정보로 구성한다. 이렇게 구성된 정보는 불필요한 정보가 감소되어 데이터의 크기가 작아지는 장점을 갖는다. 또한 구성된 정보는 주요 부분부터 단계적으로 클라이언트에 보냄으로써 클라이언트에 대한 응답시간이 빠르게 되며, 사용자의 요구에 적합한 형태의 공간 정보를 우선순위가 높은 데이터부터 단계적으로 클라이언트에서 보여짐으로써 사용자에게 필요한 공간정보를 보다 빠르게 얻을 수 있다.

1. 서 론

최근 무선 인터넷과 PDA, HPC와 같은 모바일 장치가 급속하게 발전하였고, 사용자에게 널리 보급되고 있다. 이에 따라 차량 네비게이션 시스템과 같이 모바일 장치에서 지도정보를 이용하려는 사용자가 증가하였고, 이를 위한 지도정보 서비스가 발전하였다. 모바일 장치는 데스크탑과 같은 컴퓨터와 달리 자원에 한계가 있고, 네트워크의 대역폭이 적은 특징이 있다. 따라서 서버에서 모바일 장치에 맞는 지도정보를 생성하여 전송하는 Map Adaptive Application이 등장하였다[1,2,3,4]. Map Adaptive Application은 사용자의 요청에 따른 빠른 처리를 위하여 같은 공간 정보에 대하여 축척에 따라 5레벨 정도로 나누어진 데이터를 중복하여 갖기 때문에 대용량 데이터의 효율적인 관리가 필요하며, 다수의 사용자가 빠른 주기로 공간 정보를 요청하므로 많은 사용자의 요청을 동시에 처리할 수 있어야 한다. 비공유 공간 데이터베이스 클러스터의 경우 독립적인 질의처리가 가능한 공간 데이터베이스를 고속의 네트워크를 이용하여 그룹으로 묶어 놓은 것으로, 공간 데이터에 대한 분할정책으로 대용량의 데이터를 효율적인 관리가 가능하며, 복제정책으로 가용성과 급증하는 사용자의 검색질의에 대해 병렬처리가 가능하다[5,6,7,8]. 따라서 Map Adaptive Application의 기반 시스템으로 기존의 단일 데이터베이스를 사용하는 것보다 비공유 구조의 공간 데이터베이스 클러스터를 사용하는 것이 훨씬 효율적이다.

Map Adaptive Application에서는 클라이언트의 자원과 네트워크 대역폭에 따라 서버에서 공간 정보를 알맞게 생성하여 보내게 된다. 따라서 공간 정보에 대해 크기를 줄이는 것과 공간 정보에 대한 효율적인 전송 방법이 필요하게 된다. 이를 위해 기존에는 스케일 기반의 데이터 전송 기법과 벡터 데이터의 점진적인 전송 기법이 존재하였다. 스케일 기반의 데이터 전송 기법이란 출력되는 스케일에 따라 보여질 필요가 없는 피쳐들을 필터링하고, 보여지는 피쳐만을 최종적으로 전송하는 기법으로 클라이언트에 전송해야 할 결과 데이터가 적은 장점이 있으나 클라이언트가 요청한 정보에 대해 모든 연산이 끝나 완벽한 결과가 나와야 클라이언트에 전송할 수 있으므로 사용자 응답시간이 느리게 되었다. 반면에 벡터 데이

터의 점진적인 전송 기법이란 사용자가 요청한 공간 정보에 대하여 간략화된 정보부터 전송하여 점진적으로 모든 데이터를 전송하는 기법으로 사용자 응답시간이 빠른 장점이 있으나 모든 데이터를 전송해야 하므로 전송해야 할 데이터양이 큰 단점이 있다.

본 논문에서는 비공유 데이터베이스 클러스터에서 한정된 자원에 적합한 형태로 공간데이터를 처리하는 기법을 제안한다. 이는 스케일 기반의 데이터 전송 기법과 벡터 기반의 다단계 전송 기법을 혼합한 형태의 전송 기법으로 클라이언트가 요청하는 공간 정보에 대하여 단일장치의 자원과 네트워크의 대역폭에 따라 적합한 공간 정보를 생성하고, 생성된 공간 정보에 대하여 중요 순서에 따라 단계적으로 클라이언트에게 전송한다. 따라서 불필요한 피쳐들을 필터링하여 결과 데이터의 크기를 줄이며, 단계적인 전송으로 인해 사용자 응답시간이 빠른 장점을 갖는다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문과 관련된 연구를 살펴보고, 3장에서는 본 논문의 기반이 되는 시스템 환경에 대해 살펴본다. 4장에서는 제안 기법에 대해 설명하며, 5장에서 결론을 맺는다.

2. 관련 연구

2.1 스케일 기반의 데이터 전송 기법

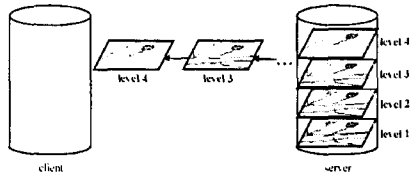
공간 정보는 특정 스케일에서 일부 피쳐만이 보여지게 된다. 따라서 스케일에 따라 사용자에게 보여지지 않는 정보가 발생하게 된다. 스케일 기반의 데이터 전송 기법은 사용자가 요청한 공간 정보에 대하여 해당 스케일에서 보여질 필요가 없는 피쳐들을 필터링하고, 보여지는 피쳐들만을 최종적으로 전송하는 기법이다. 따라서 스케일에 따른 필터링을 통해 불필요한 데이터들을 제거함으로써 결과로 전송해야 하는 공간 정보의 크기가 작아지게 되고, 이로 인해 좁은 대역폭의 네트워크에서도 빠른 전송이 가능하다. 그러나 이 기법은 클라이언트가 요청한 정보에 대하여 모두 수집하고 필터링이 끝나 완벽한 결과가 나와야만 클라이언트에 전송할 수 있으므로 사용자 응답시간이 느리게 되었다[9,10].

* 본 연구는 정보통신부의 대학 S/W 연구센터 지원사업의 연구 결과임

2.2 벡터 데이터의 점진적인 전송 기법

벡터 데이터는 클라이언트에서 필요에 따라 융통적이고 효율적으로 데이터에 대한 처리가 가능하다. 그러나 전송해야 할 데이터의 용량이 많아 한번에 모든 정보를 전송할 수 없다. 따라서 벡터 데이터의 점진적인 전송 기법에서는 해당 데이터에 대하여 상세화에 따라 5레벨 정도의 공간 정보를 두고 간략한 정보부터 차례로 모든 데이터를 보내게 된다[11,12]. 예를 들어, 고속도로에 대한 데이터를 전송하는 경우 처음에 선으로만 보이던 고속도로가 레벨이 낮아짐에 따라 이중 선으로 표시되고 다리와 인터체인지가 표시된다.

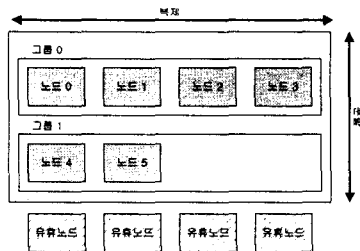
벡터 데이터의 점진적인 전송 기법은 간략화된 정보부터 전송하여 점진적으로 모든 데이터를 전송하므로 사용자 응답시간이 빠른 장점이 있으나 모든 데이터를 전송해야 하므로 전송해야 할 데이터량이 큰 단점이 있다.



[그림 1] 벡터 데이터의 점진적인 전송

3 시스템 환경

제안 기법의 기반이 되는 시스템은 비공유 공간 데이터베이스 클러스터인 GMS/Cluster이다. GMS/Cluster의 클러스터 구성은 [그림 2]와 같다.



[그림 2] GMS/Cluster의 클러스터 구성

각 노드는 독립적인 질의 처리가 가능한 GMS/Cluster 시스템으로 구성되며, 2개 이상의 노드들을 묶어 하나의 그룹을 구성한다. 그룹 내에 존재하는 각 테이블은 하나의 노드에 마스터가 존재하고 나머지 노드들에 해당 테이블의 복제본들이 존재하게 된다. 또한, 하나의 그룹에는 여러 개의 테이블이 존재할 수 있고, 각 테이블에 대한 마스터 노드의 위치는 테이블 생성시 사용자가 지정할 수 있으며, 분할된 마스터들은 그룹간 구분을 위한 고유번호를 갖는다. 분할정책은 그룹 사이에 적용되며, 복제정책은 그룹 내에 적용된다. 각 노드들은 클러스터 구성을 위해 확장된 메타정보를 유지하며, 노드간 정보 전송을 위해 고속의 네트워크로 연결되어 있다. GMS/Cluster는 급증하는 사용자 질의에 대하여 유연하게 대처하여 안정적인 서비스가 가능한 온-라인 환경을 제공하며, 이를 위해 서비스를 하지 않는 유휴 노드들이 존재한다[13].

4. 한정된 자원에 적합한 형태로의 공간데이터 처리 기법

본 장에서는 본 논문의 제안 기법의 한정된 자원에 적합한 형태로의 공간데이터 처리 기법에 대해 설명한다. 이를 위해 4.1절에서는 데이터에 대한 우선순위를 정의하고, 4.2절에서는 불필요한 데이터를 제거하고 공간데이터의 크기를 줄이기 위한 필터링 기법에 대해 살펴본다. 그리고 4.3절에서는 필터링된 공간데이터에 대하여 클라이언트에 전송하는 기법에 대해 살펴본다.

4.1 데이터 우선순위

공간 데이터에 대해 효율적인 전송을 하기 위하여 클라이언트가 요청한 데이터의 우선순위가 필요하다. 데이터의 우선순위는 먼저 공간 객체들의 분류에 따라 중요도에 따른 우선순위가 필요하며, 다음으로 같은 분류의 공간 객체들 사이에서의 우선순위가 필요하다.

공간 객체들의 분류에 따라 중요도에 따른 우선순위가 필요하며, 다음으로 같은 분류의 공간 객체들 사이에서의 우선순위가 필요하다.

[표 1]은 공간 객체 분류에 따른 중요도를 나타낸 것이다. 공간 데이터의 경우 사용자가 가장 먼저 보는 것이 도로이며, 다음으로 건물이다. 그 다음으로는 건물에 대한 건물명을 보게 된다. 이는 "중로타워에서 세종문화회관 방면으로 가면 광화문이 나온다."라는 정보와 같이 목적지를 찾아가기 위해서는 도로가 가장 필요하고 목적지를 찾아가는 길목에 있는 주요 건물들에 대한 정보가 다음으로 필요하게 된다.

[표 1] 공간객체 분류에 따른 중요도

공간객체 분류명	중요도	
	높음	낮음
도로	도로명	지하철
건물	건물명	강 산

[표 2]는 분류된 공간 객체에서 세부 객체에 대한 중요도를 나타낸 것이다. 도로의 경우 일반도로가 가장 중요도가 높으며 다음으로 자동차전용도로이다. 이는 일반적으로 고속도로나 자동차전용도로보다는 일반도로가 더 많이 사용되기 때문에 가장 중요하다. 건물의 경우에는 주요건물이 가장 중요한데, 주요건물이란 사람들이 가장 이해하기 쉬운 건물로써 학교, 공공기관, 대기업건물과 널리 알려진 건물(지하철이나 버스정류장 등에서 주변지역도에 나타나는 건물) 등이 이에 속한다. 건물명은 건물의 중요도와 같으며, 도로명의 경우 고속도로명이 가장 중요하며 다음으로 도로번호가 중요하게 된다. 이는 대부분의 길에 대한 정보를 표현할 때 "42번 국도의 어느 방면이 정제입니다."와 같이 도로번호를 자주 사용하기 때문이다. 강의 경우에는 큰 강에 대한 정보가 중요하며 다음으로는 이름이 있는 큰 하천에 대한 정보가 중요하다. 일반하천 및 개울의 경우는 지도의 상세부분을 표현하기 위한 부분으로 대부분의 사용자에게 큰 의미를 갖지 않는다.

[표 2] 분류된 공간객체들의 세부 중요도

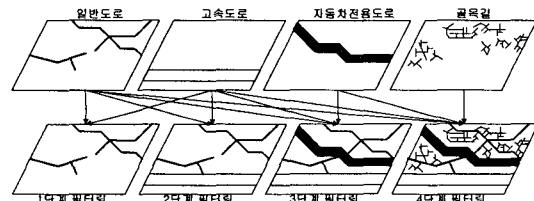
공간객체 분류명	중요도		
	높음	중간	낮음
도로	일반도로	고속도로	자동차전용도로
건물	주요건물	교차로건물	아파트 일반건물
건물명	주요건물명	교차로건물명	아파트명 일반주택명
도로명	고속도로명	도로번호	골목길명
강	강	큰하천	일반하천 및 개울

4.2 필터링 기법

필터링 기법은 불필요한 데이터를 제거하고 공간데이터의 크기를 줄이기 위한 방법으로 결과 데이터에 대한 크기를 줄이게 된다. 필터링 기법에는 공간객체 중요도에 따른 필터링과 축척에 따른 필터링을 사용한다.

4.2.1 공간객체 중요도에 따른 필터링

공간객체 중요도에 따른 필터링이란 4.1절에서 살펴본 바와 같이 공간객체들의 세부 중요도에 따라 필터링을 하는 것이다. 이는 사용자의 요청 정보에 기반을 두게 된다. 즉, 사용자는 4단계(본 논문에서는 중요도의 단계를 4개로 나누었다.) 중의 한 단계에 대한 정보를 요청한다. 서버는 사용자의 요청에 따라 해당하는 단계의 공간객체들을 선별하여 사용자에게 보내게 된다. 예를 들어 사용자가 2단계의 정보를 요청할 경우 서버는 도로에 대해서는 고속도로와 자동차전용도로만을, 건물에 대해서는 주요건물과 교차로건물만을 필터링하여 사용자에게 보내게 된다.



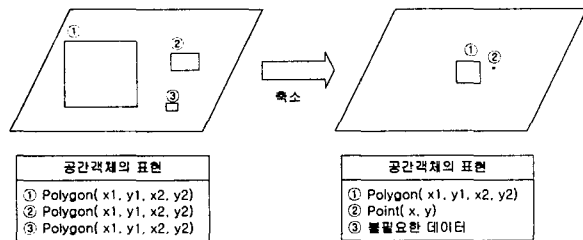
[그림 3] 공간객체(도로) 중요도에 따른 필터링 모형

[그림 3]은 도로에 대한 중요도에 따른 필터링 모형을 나타낸다. 그림에서 보면 중요도에 따라 필터링되어 표현되는 객체의 수가 다르다. 즉, 낮은 단계를 포함하는 4단계로 갈수록 더 자세한 도로가 표현된다.

4.2.2 축척에 따른 필터링

공간객체 중요도에 따른 필터링 후에도 축척에 따라 공간객체가 사용자에게 불필요하게 되어질 수 있다. 예를 들어 가로세로 1cm의 사각형으로 보이는 객체가 축척에 따라 축소되었을 경우 점으로 표시될 수 있다. 이 경우 꼭 표시되지 않아도 되는 객체의 경우에는 필터링 과정에서 제거되고, 그렇지 않은 객체의 경우에는 사각형의 객체를 점으로 변환된 정보를 갖게 함으로써 결과데이터의 크기를 줄일 수 있다.

[그림 4]는 이를 나타낸 그림으로 왼쪽의 경우 ①, ②, ③의 사각형 객체가 존재한다. 이를 공간객체로 표현하면 모두 폴리곤(Polygon)이 된다. 그러나 이를 축소할 경우 오른쪽과 같이 ①과 ②로만 표현이 되며 ③의 경우는 꼭 표시되지 않아도 되는 객체로 제거되게 된다. 이때 각각을 공간객체로 표현하면 ①은 폴리곤이 된다. 그러나 ②는 원래 사각형이지만 축소되어 점으로 보이므로 점(Point)으로 표현하게 된다.

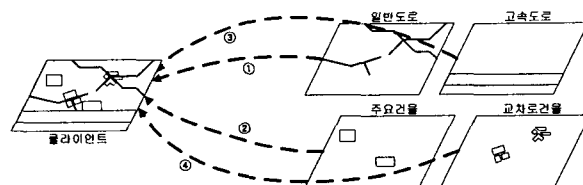


[그림 4] 축척에 따른 필터링 모형

4.3 다단계 데이터 전송 기법

필터링과정을 마친 데이터는 클라이언트에게 전송이 된다. 그러나 이때 모든 데이터의 필터링을 마친 후 전송을 하게 되면 응답시간이 느릴 뿐만 아니라 데이터의 크기도 커지게 된다. 따라서 본 논문에서는 중요도에 따라 단계별로 데이터를 전송하는 기법을 제안한다. 이는 4.1절에서 살펴본 중요도에 기반한 것으로 공간객체의 분류에 따른 중요도가 분류된 공간객체들의 세부 중요도에 우선한다. 즉, 클라이언트가 정보를 요청하면 서버는 먼저 도로의 1단계 정보를 클라이언트에 전송하고, 다음으로 건물의 1단계 정보를 전송한다. 공간객체의 분류에 따른 각 1단계의 정보를 모두 보내고 나면 다음으로 공간객체의 분류에 따른 중요도에 따라 각 공간객체의 2단계 정보를 전송한다. 이는 중요한 데이터를 먼저 전송하여 사용자에게 보여줌으로써 응답시간을 빠르게 할 뿐만 아니라 사용자가 원하는 데이터를 보다 효율적으로 전송할 수 있다.

기존의 기법의 경우 다른 레벨의 정보를 보낼 경우 해당 레벨에 포함되는 전체 데이터를 보냈으나 제안 기법의 경우 단계별로 해당 지역의 나머지 데이터들을 보냄으로써 이전에 보낸 데이터와 중복이 없어 더 적은 크기의 데이터를 보낼 수 있다. 또한, 일반도로나 골목길과 같이 중요도가 낮은 데이터의 경우 데이터의 크기가 크기 때문에 나중에 전송함으로써 사용자에게 효율적으로 공간 정보를 제공할 수 있다.



[그림 5] 다단계 데이터 전송 기법에 따른 전송

[그림 5]에서 보면 도로와 건물에 대하여 2단계 까지의 데이터를 클라이언트에 전송하는 것을 나타낸다. 이는 먼저 일반도로를 전송한 후 주요건물을 전송한다. 그 다음 고속도로를 전송한 후 교차로건물을 전송함으로써 사용자의 요청 데이터 모두에 대한 전송을 마친다. 클라이언트에서는 [그림 5]의 왼쪽 그림과 같이 받은 데이터에 대하여 순차

적으로 그려 모든 데이터가 전송된 후에는 요청한 정보에 대한 완전한 공간정보를 얻을 수 있게 된다.

4. 결론 및 향후연구

본 논문에서는 비공유 데이터베이스 클러스터에서 한정된 자원에 적합한 형태로 공간데이터를 처리하는 기법에 대하여 제안하였다. 제안 기법은 대량의 공간데이터를 효율적으로 저장 및 관리하고 수많은 사용자에게 대해서도 안정적인 서비스가 가능한 비공유 공간 데이터베이스 클러스터를 기반으로 하였다.

한정된 자원과 적은 대역폭을 갖는 네트워크에 대해 크기가 크고 복잡한 벡터 데이터를 효율적으로 전송하기 위하여 본 논문에서는 한정된 자원에 적합한 형태의 공간데이터 처리 기법을 제안하였다. 제안 기법은 클라이언트가 요청한 정보에 대하여 불필요한 데이터를 제거하고 공간데이터의 크기를 최대한 줄이는 필터링을 하고, 필터링이 끝난 데이터에 대하여 우선순위에 따라 주요 데이터부터 단계적으로 전송하였다. 따라서 결과 데이터에 대한 크기가 작고 사용자 응답시간이 빠른 장점을 갖게 되었다. 이는 자원의 한계와 적은 네트워크의 대역폭을 갖는 모바일 장치에 대하여 효율적인 공간 데이터 전송을 가능하게 하였다.

향후 연구로는 비공유 데이터베이스 클러스터에서 한정된 자원에 적합한 형태로 공간데이터를 처리하기 위해 보다 효율적인 공간 데이터의 분할 및 복제 기법이 필요하다.

5. 참고문헌

- [1] Dan Chalmers, Morris Sloman, Naranker Dulay, "Map Adaptation for Users of Mobile Systems", Proceedings of the tenth international conference on World Wide Web, 2001
- [2] Jorg Baus, Antonio Kruger Wolfgang Wahlster, "A Resource-Adaptive Mobile Navigation System", Proceedings of the 7th international conference on Intelligent user interfaces, 2002
- [3] S. Long, K. Kooper, G. D. Abowd, and Atkeson C. G., "Rapid Prototyping of Mobile Context-Aware Applications: The Cyberguide Case Study", In Proceedings of the 2nd ACM International Conference on Mobile Computing and Networking, pages 97-107, 1996
- [4] F. Hohl, U. Kubach, A. Leonhardi, K. Rothermel, and M. Schwemm., "Next century challenges: Nexus - an open global infrastructure for spatial-aware applications.", International Conference on Mobile Computing and Networking, pages 249-255, ACM/IEEE, ACM Press, 1999
- [5] mong Li Lee, Masaru Kitsuregawa, Beng Chin Ooi, Kian-Lee Tan, Anirban Mondal, "Towards Self-Tuning Data Placement in Parallel Database Systems", ACM/SIGMOD, 2000
- [6] Svein Erik Bratsberg, "Online Scaling in a Highly Available Database", Clustra, 2001
- [7] Roger Bamford, Rafiul Ahad, Angelo Pruscino, "A Scalable and Highly Available Networked Database Architecture", Proceedings of the 25th VLDB Conference, Edinburgh, Scotland, 1999
- [8] Manish Mehta, David J. DeWitt, "Data placement in shared-nothing parallel database systems", VLDB Journal, 1997
- [9] 양신, 이충호, 배해영, "인터넷을 통한 벡터 공간 데이터의 효율적 전송을 위한 최적화 기법", 한국정보과학회 논문지D, pages 273-285, 2003
- [10] Esri White Paper, "Map Generalization in GIS: Practical Solutions with Workstation ArcInfo Software", ESRI Inc., July 2000
- [11] Michela Bertolotto, Max J. Egenhofer, "Progressive Transmission of Vector Map Data over the World Wide Web", <http://www.spatial.maine.edu/~max/progressive.pdf>
- [12] Michela Bertolotto, "Progressive Vector Transmission", ACM GIS'99, 1999
- [13] 유병섭, 김영근, 김재홍, 배해영, "GMS/Cluster 설계 및 구현", 개방형 지리정보 시스템 학회, p. 225 - 231, 2003