

# PDA 지도 브라우저를 위한 동적 공간 색인 구조

김은영<sup>0</sup>, 전봉기, 서영덕, 홍봉희

부산대학교 컴퓨터공학과

{eykim2, bgjun, ydseo, bhong}@hyowon.cc.pusan.ac.kr

## A Dynamic Spatial Indexing Structure for Map browsing on PDA

Eun-Young Kim<sup>0</sup>, Bong-Gi Jun, Young-Duk Seo, Bong-Hee Hong  
Dept. of Computer Engineering, Pusan National University

### 요약

컴퓨터와 통신 기술의 발달은 휴대용 단말기의 기능을 일반 PC의 수준으로 끌어올리며 휴대 단말기의 보급을 가속화하고 있다. 이에 따라 휴대용 무선 단말기에서의 지도서비스에 대한 다양한 요구가 날이 높아지고 있다. 하지만 지도 서비스에 필수적인 기존의 공간색인은 서버나 PC와 같은 저장용량이 크고 연산처리속도가 높으며 유선망을 통해 다른 서버와 통신하는 컴퓨팅 시스템을 대상으로 하므로 PDA로 대표되는 휴대용 무선 단말기에 적용하기에는 부적합하다. 또한 휴대용 기기의 이동성을 고려할 때, 질의 수행 시 저장되지 않은 지도 데이터는 무선 통신을 이용하여 서버로부터 전송되어져야 한다. 그리고 새롭게 전송된 지도 데이터를 기존에 저장된 지도 데이터에 삽입하기 위해서 색인 재구축 비용이 발생한다. 즉, 서버와 무선통신을 하는 휴대용 단말기에서 지도서비스를 하기 위해서는 휴대 기기에 적합한 데이터 저장 및 색인구조가 필요하다.

이 논문에서는 무선단말기, 특히 PDA 환경에서의 백터지도시스템을 위해 공간 데이터의 최적화된 저장구조와, 비연속적인 다양한 지역에 대한 데이터를 효율적으로 통합·관리하는 동적색인구조를 제시한다.

### 1. 서론

휴대용 단말기 하드웨어의 빠른 발전과 IMT2000 및 UMTS(Universal Mobile Telecommunication System)으로의 초고속 무선통신의 발전은 휴대용 무선단말기의 보급을 더욱 가속화하고 있다. 특히, PDA는 전자수첩보다 더 다양하고 강력한 기능을 지원할 뿐 아니라 노트북이나 PC보다 저렴하고 이동성이 뛰어나다는 장점들로 인해 사용이 더욱 급속히 늘어나고 있다. 이에 따라, 위치기반 교통정보서비스, 물류관제, 전자지도 서비스 등 휴대용 무선 단말기에서의 다양한 지도서비스에 대한 요구 또한 높아지고 있다.

그러나 기존의 백터지도시스템은 백터지도의 높은 연산비용과 많은 데이터량으로 인해서 대규모 저장공간/메모리, 높은 CPU의 연산능력을 가진 컴퓨팅 환경을 대상으로 하고 있다. 따라서, 기존의 백터지도 데이터의 저장기법이나 색인을, PDA와 같이 비교적 자원이 적고 이동성을 가지며 무선통신을 하는 컴퓨팅 시스템에서 그대로 적용하기에는 비효율적이다. 그러므로 제한적인 자원을 가진 새로운 컴퓨팅 환경에 적합한 백터지도데이터의 저장 및 색인 구조를 제시할 필요가 있다.

이 논문에서는 이동 중인 휴대용 무선단말기에서 백터지도를 저장 및 검색하고, 서버와의 통신을 통한 지도데이터 다운로드를 지원하기 위한 효율적인 동적 공간 색인 구조를 제시한다.

이 논문의 구성은 다음과 같다. 먼저 2장에서는 관련 연구 기술

하고 3장에서는 동적 혼성 색인구조에 대해서 설명한다. 마지막으로 4장에서는 결론 및 향후 연구를 기술한다.

### 2. 관련연구

기존의 공간 색인으로는 그리드파일로 대표되는 영역기준 분할방식의 색인과 R-tree 계열로 대표되는 객체기준 분할방식의 색인이 있다. 그리드파일은 해쉬 기반의 색인으로서 고정된 도메인 공간(Domain space)에서의 색인을 구축하고, R-tree 계열의 색인은 삽입되는 객체에 따라서 도메인 공간이 가변적이다.

이동 중인 휴대 단말기에서 다양한 영역의 상대적으로 많은 데이터를 서버로부터 전송 받아 저장하게 될 경우, 휴대 기기에 저장되는 공간 데이터의 도메인 공간은 질의영역에 따라 가변적이며, 전송된 데이터는 이미 구축된 색인에 Bulk-Insertion 된다.

그리드파일은 고정된 도메인 공간을 대상으로 색인을 구축하므로, 검색 속도가 빠른 장점이 있지만 도메인 영역이 바뀌는 데이터의 삽입 시 전체색인 재구축의 비용이 든다.

[3]에서는 기 구축된 R-tree에 많은 데이터를 삽입하는 Bulk-Insertion의 방법으로써 STL(Small-Tree-Large-Tree)를 제시하였다. 그러나 이 방법 역시 작은 트리를 만들기 위한 비용이 필요하며, 그리드파일에 비해 동일한 데이터에 대해 색인을 구축하는 비용이 훨씬 크다.

3. 동적 혼성색인구조

3.1 대상 환경 및 문제 정의

이 논문은 사용자가 GIS 기능을 가진 휴대 기기(예를 들어, PDA)를 가지고 이동을 하면서 단말기에 저장되지 않은 영역의 지도데이터를 무선통신을 통해서 서버로부터 전송 받아서 저장하는 환경을 대상으로 한다(그림 1). 이러한 환경에서 질의영역은 사용자의 위치를 중심으로 하는 인근 지역일 가능성이 높다. 따라서 서버로부터 무선통신을 통해서 전송된 지도데이터는 재사용될 가능성이 높으므로 휴대단말기(이하 PDA)에 저장하고, 새로운 질의 수행 시 통신비용을 줄이기 위해서 저장되지 않은 영역의 데이터만 서버에 요청하여 전송 받는다.

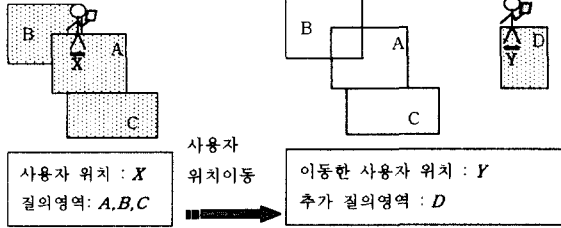


그림 1. 사용자의 위치에 따른 질의영역의 변화

위의 경우 휴대 단말기의 입장에서 다음과 같은 시나리오를 가진다.

- ① 사용자는 X 위치에서 질의 영역 A의 지도 데이터를 저장하고 있으며, A 영역 내에서 질의를 수행한다.
- ② 사용자는 A 영역에 인접한 B, C 영역에 대해 질의할 경우 저장되지 않은 영역의 지도데이터를 서버에게 요청하여 전송받아서 PDA에 저장한다. 이후, 사용자는 저장된 영역  $(A \cup B \cup C)$ 에 대해서 질의를 수행한다. B 영역에 대해  $B \cap A \& (A \cap B) \neq \emptyset$ 의 조건을 만족하므로 (B-A) 영역의 지도데이터를 전송받는다.
- ③ 사용자가 위치 X와 동떨어진 D 영역 내의 위치 Y로 이동한다. 이동한 위치 Y에서 D 영역에 대한 데이터를 무선통신으로 서버로부터 전송받아서 저장한다.

그리드파일로 대표되는 해쉬 기반 공간색인은 고정된 도메인 공간에서 검색속도가 비교적 일정하고 빠르며, 초기 색인 구축 시 한 번의 데이터 스캔으로 색인을 구축할 수 있다는 장점이 있다[4]. 그러나 위의 예시와 같이 저장되지 않은 질의영역의 지도데이터를 서버로부터 전송받아 저장할 때마다 도메인 공간이 아래 그림 2와 같이 A,  $MBR(A \cup B)$ ,  $MBR(A \cup B \cup C)$ ,  $MBR(A \cup B \cup C \cup D)$ 로 확장된다. 따라서 이 논문의 대상 환경에서 그리드파일로 색인을 구축할 경우 도메인 공간이 증가할 때마다 전체 색인을 재구축 하는 문제가 발생하게 된다. 또는 서버 측의 도메인 공간을 클라이언트(PDA) 도메인 공간으로 설정하고 그리드파일을 구축하면 PDA의 제한적인 저장용량을 고려할 때 색인의 용량이 커져서 공간 활용도가 떨어진다는 문제점이 있다.

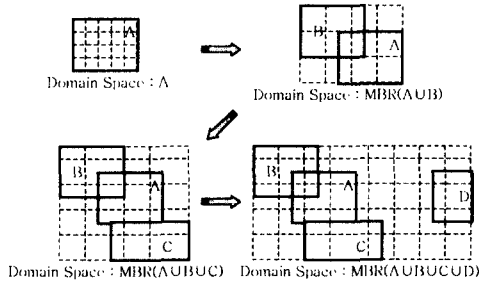


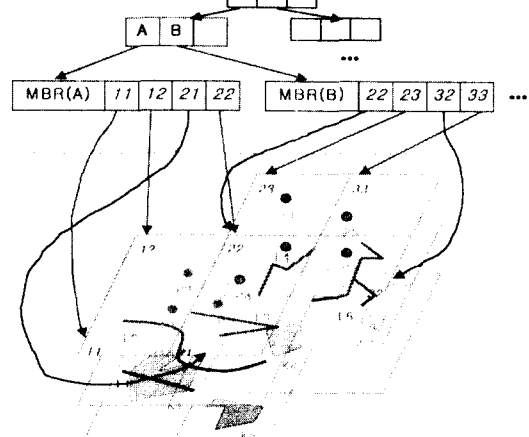
그림 2. 도메인 공간의 확장

그리드파일과는 달리 R-tree는 도메인 공간이 미리 고정된 것이 아니라 삽입되는 객체에 따라 달라질 수 있다. 그러나 R-tree에 많은 객체를 한번에 삽입(Bulk-Insertion)하는 비용은 그리드파일에 비해 크기 때문에 [3] 사용자의 질의에 대해 빠른 응답시간을 보장할 수 없다. 그리고 기존 색인만으로는 PDA에서 저장하고 있는 지도 영역에 대한 정보를 관리할 수가 없기 때문에, 사용자가 질의를 했을 때 이미 저장된 영역과 서버에서 추가적으로 전송 받을 영역을 구분할 수 없는 문제점이 있다.

3.2 동적 혼성색인구조

이 논문에서 제시하는 동적 혼성색인구조는 아래 그림 3과 같다. 그림 3에서 동적 혼성색인구조는 (1)질의영역 색인(QueriedRegion-Index)과 (2)캐쉬데이터 색인(CachedData-Index)으로 구성된다.

(1) 질의영역 색인



(2) 캐쉬데이터 색인

그림 3. 동적 혼성색인구조

질의영역 색인은 질의영역들에 대한 색인으로서, 도메인 공간의 층간이 자유로운 트리 계열의 색인을 사용한다. 질의영역에 대한 정보는 질의영역의 MBR 정보와 캐쉬데이터 색인인 가상고정그리드파일에서의 해당 셀 아이디(Cell\_ID) 및 디렉토리에 대한 pointer를 포함하게 된다. 즉, 질의영역 색인은  $[MBR(QueriedRegion), Cell\_ID\_List \& pointers]$ 에 대한 색인이다.

캐쉬데이터 색인은 각 질의영역(Q)에 포함되는 객체들에 대한 색인으로서, 고정된 도메인 공간에 대해서 빠른 구축과 검색을 지원하는 고정그리드파일을 응용한 가상고정그리드파일을 사용한다.

가상고정그리드파일은 고정된 도메인 공간을 같은 크기의 셀로 분할한다는 점에서 기존의 고정그리드파일과 같지만, 모든 셀에 대한 디렉토리를 가지는 것이 아니라, PDA가 데이터를 저장하고 있는 영역의 셀에 대해서만 디렉토리를 가진다. 따라서, 가상고정그리드파일의 디렉토리 구조는 각 셀에 대한 정보가 셀아이디에 따라 정렬된 연결리스트(Linked List)의 자료구조를 가진다.

3.3 동적 혼성색인구조를 위한 알고리즘

이 논문의 대상 환경에서 이동 중인 사용자가 영역 질의(Window Query)를 할 때, 질의 영역(Window Region, WR)은 기존에 질의했던 영역(Queried Region, QR)과 일부 겹치고  $WR_Q : WR \cap (\sum QR)$ , 일부는 겹치지 않는  $WR_S : WR - (\sum QR)$  경우가 일반적이다. 이 경우의 검색 및 삽입 알고리즘은 다음과 같다.

가. 검색 알고리즘

동적 혼성색인구조에서 검색은 크게 두 단계의 여과 과정과 한 단계의 정제 과정을 거쳐서 수행된다.

- ① **여과 과정 1** - 질의영역 색인에서의 여과 과정으로써, 새로운 질의영역(Window Region)을  $WR$ , 기존 질의영역(Queried Region)을  $QR$ 이라고 할 때,  $Overlaps(WR, QR)=True$ 를 만족하는  $QR$ 을 점진영역집합( $S\_QR$ )에 추가한다.  $S\_QR$ 이 공집합이 아닐 경우,  $S\_QR$ 의 각  $QR(MBR(QR), Cell\_ID\_list)$ 에 대해서  $Cell\_ID\_List$ 에서 실제  $WR$ 과 겹치는 셀의 아이디를 점진영역셀집합( $S\_CellID$ )에 저장하고,  $WR-(S\_QR)$  영역은 기존 질의영역과 겹치지 않는 영역이므로 서버에 질의하여 다운로드 받는다.
- ② **여과 과정 2** - 캐쉬데이터 색인에서의 여과 과정으로써, 여과 과정 1에서 만들어진 점진영역셀집합에서 각 셀에 해당하는 캐쉬데이터 색인의 디렉토리에 대한 포인터를 얻는다. 캐쉬데이터 색인의 디렉토리에서 객체들의 군사값과  $WR$ 을 비교하여 후보객체집합을 선별한다.
- ③ **정제 과정** - 여과 과정 2를 거쳐서 선별된 객체 집합에 대해서 검색조건( $WR$ )과 완전히 일치하는 객체만을 선택한다.

아래 그림 4는 사용자 가 질의를 시작한 순간부터 결과를 받을 때까지의 검색 시나리오이다.

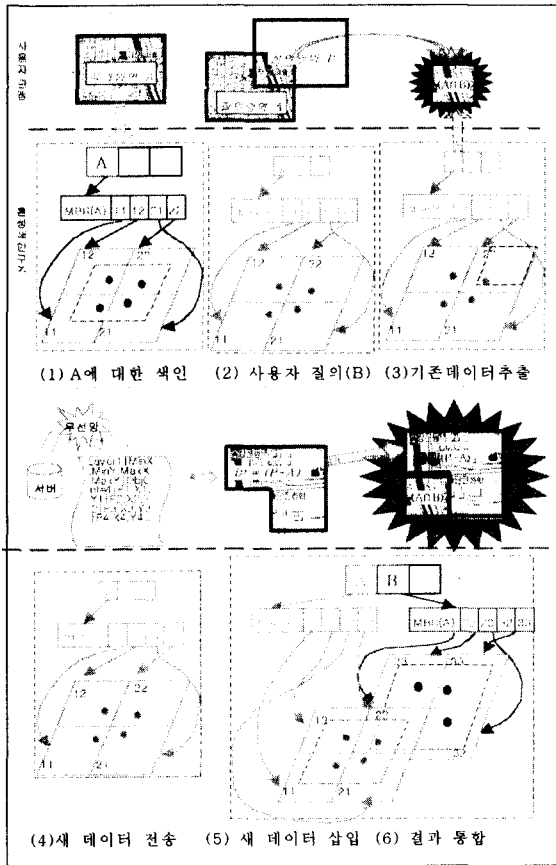


그림 4. 검색 시나리오

나. 삽입 알고리즘

이 논문의 대상이 되는 환경에서 동적 혼성색인구조에 대한 삽입은 객체 한 개의 삽입이 아니라 질의영역에 포함되는 모든 객체의 삽입, 즉 Bulk-Insertion이 된다. 즉, 위 그림 4. 검색 시나리오와 같이 질의 처리 시 저장되어 있지 않은 영역은 (4)-(5) 단계처럼 서버로부터 전송 받아서 동적 혼성색인구조에 삽입/저장하고 동시에 질의 결과로 화면에 출력해야 한다. 동적 혼성색인구조에서 삽입 알고리즘은 크게 세 단계로 나뉜다.

- ① **캐쉬데이터 색인에 객체 삽입** - 삽입될 모든 객체에 대해서 객체의 MBR이 캐쉬데이터 색인의 어느 셀에 해당되는지 셀아이디를 구한다. 셀아이디가 점진영역셀집합에 속한다면 캐쉬데이터 색인의 해당 셀에 객체를 삽입한다. 만약, 셀아이디가 점진영역셀집합에 속하지 않는다면, 캐쉬데이터 색인에 새로운 셀을 추가하고, 새로 추가된 셀에 객체를 삽입한다.
- ② **질의영역정보 수집** -  $WR$ 에 대한 MBR 정보와 1에서 객체가 삽입된 모든 셀아이디를 셀아이디리스트( $Cell\_ID\_List$ )에 추가하여 질의영역 색인에 삽입될 질의영역정보를 구성한다.
- ③ **질의영역 색인에 질의영역정보 삽입** - 2에서 만들어진 질의영역정보를 질의영역색인에 삽입한다.

4. 결론 및 향후 연구

이 논문에서는 메모리와 저장공간이 비교적 적으면서 이동성을 가지고 무선통신을 하는 휴대용 단말기에 대해 최적화된 공간데이터의 동적 혼성색인구조를 제시하였다.

동적 혼성색인구조는 비연속적인 여러 개의 질의영역에 대한 객체를 저장할 때 적은 비용으로 전체 영역에 대한 색인을 구축할 수 있다. 또한 사용자가 질의를 했을 때, 질의 영역이 현재 저장된 영역과 중복되는지 빠르게 검사하여 중복되지 않은 영역에 대해서만 서버에 질의함으로써 대역폭이 좁은 무선망에서의 전송비용을 줄일 수 있다.

향후 연구로 서버로부터 다운로드할 지도데이터가 PDA의 잔여 저장용량을 초과할 경우를 해결하기 위한 효율적인 지도데이터 교체 알고리즘이 필요하다.

5. 참고 문헌

- [1] Norbert Beckmann, Hans-Peter Kriegel, "The R-tree: An Efficient and Robust Access Method for Points and Rectangles", Proceedings of the 1990 ACM SIGMOD, pp. 332-331, 1990.
- [2] Christos Faloutsos, Timos Sellis, Nick Roussopoulos, "Analysis of Object Oriented Spatial Access Methods", Proceedings of the 1987 ACM TODS, pp. 426-439, 1987.
- [3] Li Chen, Rupesh Choubey, Elke A. Rundensteiner, "Bulk-Insertions into R-Trees", ACM-GIS 1998, pp. 161-162, 1998.
- [4] J. Niervegelt, Hans Hinterberger, Kenneth C. Sevcik, "The Grid File: An Adaptable, Symmetric Multikey File Structure", ACM TODS, pp. 38-71, 1984.
- [5] 김은영, 전봉기, 서영덕, 홍봉희, "PDA에서 공간데이터를 저장하기 위한 혼성색인구조", 2001 한국정보과학회 논문집, vol.28 no.1, pp. 34-36, 2001.