

# 위치 기반 서비스에서 비연결 연산을 위한 Layer와 Region을 고려한 의미론적 캐싱 기법

<sup>0</sup>이상철\*, 이충호\*, 김재홍\*\*, 배해영\*

\*인하대학교 전자계산공학과

\*\*영동대학교 컴퓨터공학과

philmap@korea.com

## Semantic Caching Method that consider Layer and Region for Disconnected Operation in Location Based Service

<sup>0</sup>Sang-Cheol Lee\*, Chung-Ho Lee\*, Jae-Hong Kim \*\*, Hae-Young Bae\*

\*Dept. of Computer Science & Engineering, Inha University

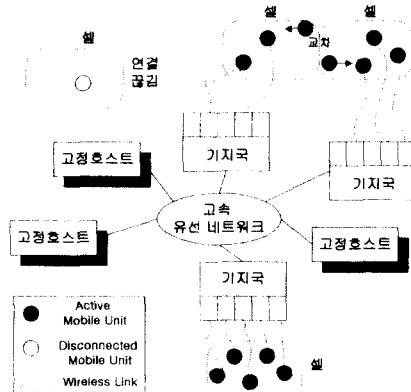
\*\*Dept. of Computer Engineering, Youngdong University

### 요약

현재 무선단말기 보급의 확대와 시간과 공간의 제약을 뛰어넘는 장점으로 무선 인터넷 환경이 급속도로 발전하고 있다. 그러나, 본 논문의 연구 분야인 모바일 지리정보 시스템(Mobile Geographic Information System)에서 무선인터넷을 통하여 실시간 지리 정보를 얻기 위해 사용자의 무선단말기와 서버 사이에 빈번한 데이터 송수신이 이뤄져야 하며 데이터의 빠른 변화를 수용해야 한다. 그러나, 아직 현저히 낮은 대역폭을 가진 채널을 통해 통신하며, 비싼 이용 요금과 오류로 인한 자발적 또는 비자발적 연결 끊김 현상 등은 서비스 제공에 한계를 가진다. 그러므로 본 논문에서는 이동 컴퓨팅 환경의 연구를 통해 연결 끊김 현상에서도 클라이언트의 지속적인 서비스 운영을 위한 비연결 연산으로 의미론적 캐싱 기법과 캐시 교체 전략에 대해 알아보고, 본 논문에서는 Layer와 Region을 고려한 의미론적 캐싱 기법을 제안하여 무선 환경에서 동적인 지도서비스를 위한 기법을 제시하였다. 이 연구를 통해 아직 많은 한계를 지닌 무선인터넷 환경에서 지리정보 뿐 아니라 위치기반의 다양한 서비스 분야에 응용될 수 있으리라 기대된다.

### 1. 서론

이동 컴퓨팅 환경은 아래의 그림1에서 처럼 고속의 유선 네트워크를 통해 고정호스트와 기지국은 상호 연결된다. 이동단말(Mobile Unit)과 기지국은 유선 네트워크보다 현저히 낮은 대역폭을 가진 무선 채널을 통해 통신한다. 이 이동단말은 제한된 대역폭에 의해 한정된 지리적 영역 내에서 자유롭게 움직이며, 이 이동성을 관리하기 위해 전체 지리적 이동 영역을 셀이라 부르는 좀더 작은 영역들로 나뉜다[1].



[그림1] 이동 컴퓨팅 모델

무선 통신에서 통신 비용의 절약을 위해 사용자가 강제로 연결을 끊을 수 있고 또는 오류에 의해 연결 끊김(Disconnection)이 종종 발생하기도 한다. 이런 현상은 데이터의 복사 없이는, 데이터의 일관성(Consistency)에 위협을 줄 수 있다. 그러나 데이터베이스 시스템은 또한 예상치 못한 연결 끊김에 있어 데이터 손실을 피하기 위해 트랜잭션 회복(Transaction recovery)의

의무를 가지고 있다. 네트워크 분할의 빈도가 높으면 높을수록 고정된 네트워크보다 더욱 강력한 에러 회복이 요구된다[6].

그러므로 본 논문에서는 이동 컴퓨팅 환경의 연구를 통해 연결 끊김 현상에서도 클라이언트의 지속적인 서비스 운영을 위한 비연결 연산으로 의미론적 캐싱 기법과 캐시 교체 전략에 대해 알아보고, 본 논문에서는 위치기반의 의미론적 캐싱 기법과 위치관련 프레디킷을 앞으로 우선 배치, 수행하는 위치기반 캐시 교체 정책을 제안하여 무선 환경에서 동적인 지도서비스를 위한 기법을 제시하였다.

본문의 구성은 2장에서 비연결 연산 기법, 의미론적 캐싱, Manhattan Distance에 대한 관련 연구를 하며, 3장에서 위치 기반 모바일 애플리케이션과 데이터에 대한 연구를 진행하며, 4장에서는 Layer와 Region을 고려한 의미론적 캐싱 기법을 제시하고, 5장인 결론에서는 Layer와 Region을 고려한 의미론적 캐싱 기법의 의의와 향후 진행될 연구에 대해 논의하였다.

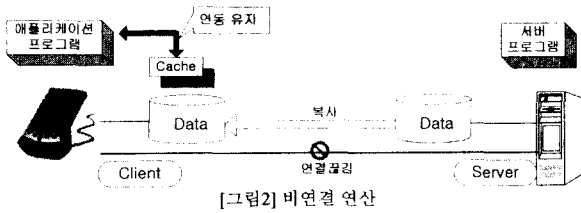
### 2. 관련연구

#### 2.1 비연결 연산(Disconnected Operation)

배터리의 한계를 가진 모바일 클라이언트는 무선 네트워크 상에서 서버와 메시지의 송수신이 이루어지게 되는데, 본 논문에서 제시할 위치기반 데이터의 경우 분산 애플리케이션에서 사용자의 장시간 데이터베이스 시스템 조작이 행해지므로 통신시 연결 끊김이 빈번히 발생할 수 있다. 연결이 끊김에 따라 연동도 단절되면 애플리케이션을 실행할 수 없게 된다. 그러므로 연결 끊김 상태에서도 연동을 유지하여 모바일 클라이언트에서의 애플리케이션 프로그램을 계속 실행할 수 있어야 한다. 이와 같이 연결이 단절되어도 애플리케이션 프로세스를 계속 동작시키는 것을 비연결 연산(Disconnected Operation)이라 한다.

비연결 연산을 처리하는 기술로 서버 내의 애플리케이션 프로세스가 이용하는 데이터를 모바일 클라이언트에 복사본을 두는 방법이 있는데 이것을 캐싱(Caching)이라 한다. 즉, 클라이언트는 서버와의 통신 연결이 끊어진 상태에서도 서버와 연결 레벨의 통신을 하지 않고 캐시된 데이터를 조작해 연동을 유지한다[7].

\* 본 연구는 정보통신부의 대학 S/W 연구 센터 지원사업의 결과임



[그림2] 비연결 연산

2.2 의미론적 캐싱(Semantic Caching)

페이지 또는 튜플 캐싱 등의 전통적인 캐시 구조는 캐시된 데이터에 대한 의미론적 정보(Semantic Information)를 갖고 있지 않다. 그러므로 내용기반의 추론을 지원할 수 없게 된다. 그러나, 의미론적 캐싱은 이동 클라이언트가 캐시내에 의미론적 기술과 이전 질의에 관련된 응답결과를 유지하므로 새로운 질의가 발생시 이동 클라이언트가 캐시내에서 전체적으로 응답할 수 있는지를 결정하게 되며, 얼마나 많은 부분을 응답할 수 있는지를, 그리고 이동단말기기의 저장용량의 한계에 따라 적용할 캐시 교체 정책으로 어떤 데이터를 제거 할 것인가를 결정 할 수 있게 한다[5]. 이를 통해 이동중인 사용자의 질의에서 이전에 캐시된 부분과 캐시되지 않은 부분을 분리하는 질의분석 과정을 용이하게 하여 이전에 캐시된 부분은 로컬에서 응답하며, 캐시되지 않은 부분은 서버를 통해 데이터를 전송 받게 된다.

2.3 Manhattan Distance

Manhattan distance 는 이동성 기반의 의미론적 캐시 교체 정책에서 적용된 기법으로, directional Manhattan distance function 을 사용하여 semantic value 를 계산하여 Victim 을 선정하게 된다. 즉, 사용자의 현 위치로부터 Manhattan distance 기반의 각 캐시된 의미론적 영역을 위한 value 를 계산 후, 계산된 values 에 따라 영역(region)들을 버린다[3].

그러나, 이 기법은 이동중인 사용자가 던진 질의에는 캐시된 부분과 캐시되지 않은 부분이 발생시에 중첩되는 질의 영역과 중첩되지 않는 2개 부분으로 분리되는데 이때 Manhattan distance 를 계산하기 위하여 각각의 분리된 영역의 중심을 계산하는데 어려움이 있다. 그러므로 본 논문에서는 의미론적 캐시 부분에 위치정보를 고려하게 되며, 위치기반 프리디킷을 우선 재배치하려 한다.

3. 위치 기반의 모바일 애플리케이션과 데이터

Public data 는 수직적 애플리케이션에서 주로 관리되는 반면, Shared data 는 수평적 애플리케이션에서 사용되며 약간의 중복된 복사본이 가능하다. 본 논문에서 연구하려는 위치 기반의 모바일 지리정보 애플리케이션은 비연결 연산이 가능한 캐싱을 위해 복사본이 가능한 Shared data 를 처리하는 수평적 애플리케이션이 되어야 한다. 이 공유된 데이터의 복사본들은 기지국과 이동국 모두에서 저장될 수 있는데, 이때 트랜잭션 관리에서 일관성 및 아키텍처의 확장성 등에서 다양한 문제들이 발생된다. 그러므로 다음과 같은 가정의 질의에서 위치 기반 데이터의 테이블을 제시하고자 한다.

• 가정

한 여행자가 차량을 가지고 이동 중에 현재 위치인 인천에서 부산으로 여행하려 할 때, 각 지역의 호텔의 위치정보, 실시간의 객실 이용 가능상황 등을 알고자 할 경우의 질의

• 응답

제한된 bandwidth의 무선통신의 상황뿐 아니라 실시간 정보의 제공이 필요한 환경에서 기존의 표1과 같이 여러 지역의 데이터를 종합적으로 관리하기 보다는 표2와3과 같이 인천에 대한 정보는 인천의 지리영역에서 부산에 대한 정보는 부산의 지리영역에서 관리하는 위치에 근거한 위치 기반 데이터의 사용은 테이블 크기의 감소, 질의의 효율성 향상, 그리고 새로운 호텔 발생으로 인한 데이터 추가 시 확장성과 지속성 등의 장점을 지닌다.

Sta	Location	Hname	Avail
10	Incheon	Songdo	Yes
6	Incheon	Olympus	No
3	Incheon	Galaxy	Yes
9	Incheon	Royal	Yes
12	Incheon	Chosun	No
3	Incheon	Paradise	No
10	Incheon	Marrtrot	Yes
5	Incheon	Heeundae	Yes

[표 1] Total Data

Sta	Location	Hname	Avail	Sta	Location	Hname	Avail
10	Incheon	Songdo	Yes	12	Pusan	Chosun	No
6	Incheon	Olympus	No	3	Pusan	Paradise	No
3	Incheon	Galaxy	Yes	10	Pusan	Marrtrot	Yes
9	Incheon	Royal	Yes	5	Pusan	Heeundae	Yes

[표 2]Incheon Data

[표 3]Pusan Data

4. Layer와 Region을 고려한 의미론적 캐싱 기법

본 장에서는 이동 호스트들에서 운영되는 애플리케이션에 따라 서로 다른 데이터의 요구 사항들을 가지게 되는데, 이런 모바일 애플리케이션과 데이

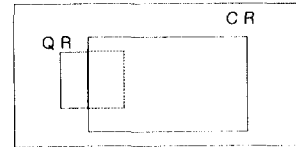
터의 상호관계를 살피고, 본 GEO/WEB시스템에서 확장된 Mobile GIS 개발을 위한 Layer와 Region을 고려한 의미론적 캐싱 기법과 교체 전략에 대한 연구를 진행하였다.

4.1 위치기반 질의 처리

이동 단말을 가진 사용자가 셀 경계들을 통과하거나 건너갈 경우 정확한 질의 결과들을 받기 위해 데이터의 위치와 방향성을 파악하는 것이 중요하다. 예를 들어, 자동차를 가지고 인천에서 부산으로 가려는 사용자가 자신이 가는 지역의 호텔의 위치정보 및 숙박 이용 가능성을 파악하려 할 때, 이동 클라이언트는 방향성을 가지며, 주기적인 서버와의 통신을 통해 해당지역의 좌표값을 얻어내고, 클라이언트의 캐시에 데이터의 수정된 시간을 타임 스탬프(TS:Time Stamp)로 나타내어 데이터의 유효성을 결정하게 된다. 즉, 여러 상황으로 자발적 또는 비자발적 연결 끊김 현상이 발생시 클라이언트는 가장 최근에 타임 스탬프 유�효성이 표시된 캐시 데이터를 사용하여 연결 끊김 현상을 극복하게 된다.

위치기반 질의 처리의 단계는 다음과 같다.

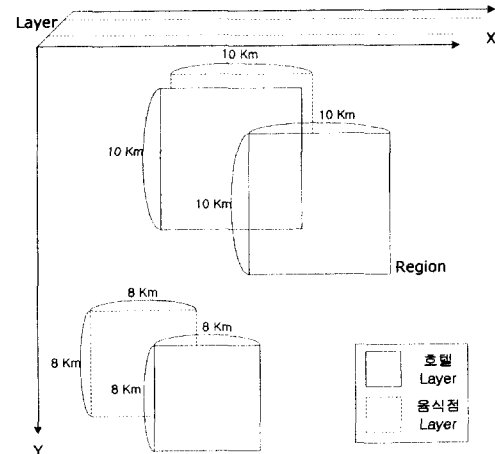
즉, 첫 단계에서 상용하는 프리디킷 위치 바인딩을 수행한다. 여기서 바인딩이란 Mobile Unit으로 질의를 보낸 사용자의 위치를 최종적으로 얻어내는 것으로 한 쌍의 위도,경도의 좌표를 얻게 된다. 본 논문은 최초 질의 요청을 받았을 때의 이동 단말기의 위치를 바인딩한다고 가정한다. 두 번째, 위치기반 질의는 일반적인 데이터베이스 질의로 변환된다. 세 번째, 현재의 위치에서 질의분석을 통해 질의에서 요구된 영역(QR)이 그림3과 같이 캐시된 영역(CR)에 일부 교차된 경우 질의 분리가 이뤄진다. 그 후 던져진 질의중 캐시된 부분과 교차하는 영역은 우선 로컬 캐시로부터 처리하며, 캐시되지 않은 부분은 서버와의 통신을 통해 데이터를 전송 받게 된다.



[그림3] 질의 분리(Query Trimming)

4.2 Layer와 Region을 고려한 Cache 알고리즘

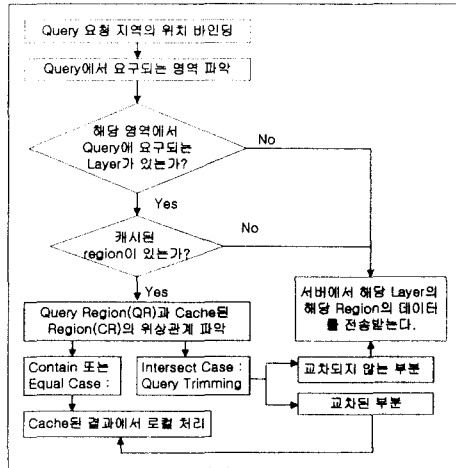
기존의 웹 지리정보시스템인 GEO/WEB에서는 사용자의 대기시간을 줄이기 위해, 전체 지리 데이터를 각각의 주제에 따라 레이어 단위로 전송하는 기법을 사용해왔다. 그러나, 해당 레이어의 일부 영역(region)을 요구하여도 항상 레이어 전체가 서버로부터 전송되어야 하는 단점으로, 레이어 전체 영역을 일정한 크기의 작은 영역인 타일(tile)로 분할,관리하여 일부 영역의 지리데이터만 요구시 해당 영역을 포함하는 부분의 지리데이터만 전송하는 기법을 적용하였었다.



[그림4] Map의 Layer와 Region에 따른 Cache

본 논문에서는 기존의 GEO/WEB 시스템을 Mobile GIS로 확장하는 가운데, 그림4와 같은 개념도를 적용하였다. 즉, 무선환경 하에서 트래픽(traffic)을 감소 시키고자 레이어 단위에 정보를 제공하면서 타일단위의 영역 캐시를 고려하였다. 그러므로, 캐시를 적용함에 있어 그림4와 같이 x축에는 해당영역의 x좌표를, y축에는 해당영역의 y좌표를, z축에는 지도의 레이어들을 나타내는 3차원 좌표축을 적용하여 동일 영역에 호텔 레이어와 음식점 레이어가 동일 지역에 중첩되었을 지라도 캐시된 영역은 없는 것으로 가정한다. 다만 동일 레이어에 대해 교차되는 영역을 캐시에서 고려하게 된다.

레이어(Layer)와 영역(Region)을 고려한 캐시 알고리즘을 적용시 질의가 요청되었던 지역의 위치 바인딩을 우선 처리하여 자신이 요청한 정보가 로컬에 캐시 되어 있는지 파악하는 것이 우선되어야 하며, 이때에 해당 레이어와 영역이 캐시 정보에 있는가의 존재유무를 파악하여 다음의 그림5와 같이 처리하게 된다.



[그림5] Map의 Layer와 Region에 따른 Cache 알고리즘 순서도

이때에 모바일 클라이언트 단말기에 캐시되는 세그먼트들을 다음의 표4와 같은 의미론적 구체적인 정보로 저장하여, 캐시 상태나 캐시 교체전략에 논리적으로 사용하게 된다.

No	Relation	Attribute	Predicate	Coord	TS
1	Hotel	Hname, Avail, Price	$(Lx-5sxsLx+5) \wedge (Ly-5sysLy+5) \wedge (Avail=Yes) \wedge (Price < 50,000)$	35,82	T1
2	Rest.	Rname,	$(Lx-3sxsLx+3) \wedge (Ly-3sysLy+3)$	43,74	T2
3	Hotel	Hname, Avail, Price	$(Lx-10sxsLx+10) \wedge (Ly-10sysLy+10) \wedge (Avail=Yes) \wedge (Price < 80,000)$	45,63	T3
4	Freeway	Fname,	$(Lx-10sxsLx+10) \wedge (Ly-10sysLy+10) \wedge (traffic < 30)$	63,50	T4
5	Hotel	Hname, Avail, Price	$(Lx-5sxsLx+5) \wedge (Ly-5sysLy+5) \wedge (Avail=Yes) \wedge (Price < 50,000)$	71,43	T5

[표4] Layer와 Region을 고려한 의미론적 캐시 인덱스의 예

#### 4.3 위치기반 캐시 관리

이동 단말기가 가질 수 있는 저장장치의 크기는 제한적이다. 하지만 실시간 지리 정보를 얻기 위해서 서버와 빈번한 데이터 송수신이 이뤄져야 하며 데이터의 빠른 변화를 수용하는 과정에서 캐시된 자료의 양은 급격히 증가하게 된다. 그러므로, 이런 현상으로 인해 이동 단말기에 캐시된 데이터가 가득찬 경우, 이전에 받아서 운영하던 캐시 데이터 중 일부를 제거하고 새로운 데이터들을 저장할 필요가 있다. 이때 어떤 데이터를 지우는 것이 적절한가를 결정하는 과정을 캐시 교체 전략이라 하며, 그 알고리즘에는 가장 오래된 문서를 지우는 FIFO방식, 가장 오래 전에 사용된 문서를 지우는 LRU방식, 잘 쓰이지 않는 문서를 지우는 LFU방식 등이 있다.

위치기반의 의미론적 캐시 기법에서는 의미론적 캐시를 의미론적 메트릭스에 의해 데이터를 구성하며, 다음에 제시할 위치기반 프레디킷 우선배치 기법을 적용함에 있어 캐시 관리를 더욱 유연성 있게 한다.

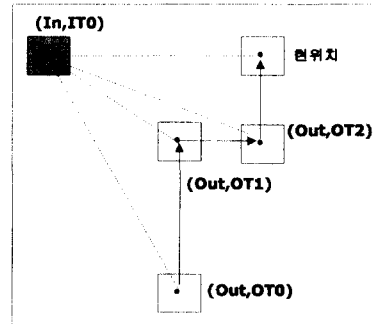
#### 4.4 위치기반 프레디킷 우선배치 기법

본 논문에서 고려하고 있는 실시간 모바일 지리정보시스템에서, 위치기반 질의 패턴은 무선 단말기를 가진 사용자의 이동성과 밀접하게 관련된다. 이동 단말기를 보유한 사용자가 위치관련 정보를 얻기 위해 새로운 질의를 던졌을 경우 이전 질의의 결과 즉, 캐시된 부분과 캐시되지 않은 부분을 분리하는데 신속히 처리하는 것이 중요하다. 그러므로 정리되어 있지 않은 프레디킷들을 표5와 같이 위치관련 프레디킷을 앞으로 우선 배치하여 저장해야 한다.

No	Predicate
1	$(Lx-5sxsLx+5) \wedge (Ly-5sysLy+5) \wedge (Avail=Yes) \wedge (Price < 50,000)$
2	$(Lx-3sxsLx+3) \wedge (Ly-3sysLy+3)$
3	$(Lx-10sxsLx+10) \wedge (Ly-10sysLy+10) \wedge (Avail=Yes) \wedge (Price < 80,000)$

[표5] 위치관련 프레디킷 우선 배치

즉, 위치관련 프레디킷을 우선 고려하여 그림6에서와 같이 region의 중심점들을 계산 후, distance를 계산하며 현위치에서 캐시된 질의 결과들을 OUT 집합에 모아서 그 중 가장 멀리 있는 캐시를 제거하는 방식을 선택한다. 제거될 캐시인 Victim선정을 신속히 처리하여 이동성을 가진 질의 처리에 반응속도를 높여 준다.



[그림6] Victim 선정 과정

#### 5. 결론 및 향후 연구

본 논문에서는 이동 컴퓨팅 환경의 연구를 통해 연결 끊김 현상에서도 클라이언트의 지속적인 서비스 운영을 위한 비연결 연산으로 의미론적 캐시 기법과 캐시 교체 전략에 대해 알아보았으며, Layer와 Region을 고려한 의미론적 캐싱 기법을 제안하여 무선 환경에서 동적인 지도서비스를 위한 기법을 제시하였다. 이 연구를 통해 아직 많은 한계를 지닌 무선인터넷 환경에서 지리정보뿐 아니라 위치기반의 다양한 서비스 분야에 응용될 수 있으리라 기대된다.

향후 연구 과제로는 본 논문에서 제시하였던 Layer와 Region을 고려한 의미론적 캐싱 기법에 대한 성능평가가 요구되며, 더 나아가 위치기반의 좀더 효율적 캐시 클러스터링 기법과 캐시 교체 전략에 대한 연구가 요구된다.

#### 참고 문헌

- [1] ELMASRI, NAVATHE "Mobile Databases", Fundamentals of Database Systems, Third Edition, ADDISON-WESLEY, 2000, pp.886-891
- [2] Daniel Barbara "Mobile Computing and Databases - A Survey", IEEE, 1999, pp.108-117
- [3] S.Dar, M.J.Franklin, B.T.Jonsson, D.Srivastava, M.Tan "Semantic Data Caching and Replacement", VLDB, 1966, pp.330-341
- [4] Margaret H.D, Vijay Kumar "Location Dependent Data and its Management in Mobile Databases", DEXA Workshop, 1998
- [5] Qun Ren, Margaret H.D "Using Semantic Caching to Manage Location Dependent Data in Mobile Computing", MOBICOM 2000, pp.210-221
- [6] Astrid Lubinski. "Security Issues In Mobile Database Access", KLUWER ACADEMIC PUBLISHERS, 1999
- [7] 일본 멀티미디어 통신연구회 "모바일 터미널시스템의 소프트웨어 기술", 그림으로 보는 모바일 컴퓨팅, 교보문고, 2000, pp.219-297
- [8] 김경배 "실시간 이동 트래픽선을 위한 동시성 제어 기법", 박사학위논문, 인하대학교, 2000
- [9] 최현호 "웹 기반 지리정보시스템을 위한 지리 데이터 분할 저장 관리 기법", 석사학위논문, 인하대학교, 1999