The Journal of GIS Association of Korea, Vol. 11, No. 3, pp. 261-274, November 2003

웹 지리정보시스템에서 다단계 추상화 데이터의 확률기반 프리페칭 기법

황병연* · 박연원** · 김유성***

Probability-based Pre-fetching Method for Multi-level Abstracted Data in Web GIS

Byung-Yeon Hwang* · Yeon-Won Park** · Yoo-Sung Kim**

요 약

기존의 웹 지리정보시스템에서는 타일링 개념을 도입하여 타일 단위로 전송하는 타일 기반의 데이터 구조를 가진다. 이는 지도의 초기 로딩 시간을 줄이는 효과는 있으나, 웹 지리정보시스템에서 사용자의 질의에 대한 전체 응답시간을 줄이지는 못하였다. 이러한 문제를 해결하기 위해 제안된 것이 프리페칭 알고리즘과 이와 연동하는 캐쉬대체 알고리즘이다.

프리페칭 알고리즘은 확률기반모델로 표현된 사용자의 타일접근패턴을 이용하여 앞으로 사용될 타일을 정확하게 예측하여 클라이언트에 미리 데이터를 전송하는 방법이고, 캐쉬대체 알고리즘은 클라이언트의 제한된 캐쉬공간에 서버로부터 가져온 타일을 캐쉬하기 위해서 사용자의 타일접근패턴을 이용하여 제거해야할 타일을 선정하는 정책이다. 웹 지리정보시스템은 확대 질의나 축소 질의 처리 시에 빠른 응답시간으로 서비스하기 위해 레벨링 기법이 사용된 다단계 추상화 데이터로 구성되어 있다. 그러나, 기존에 제안된 프리페칭 알고리즘은 다수 개의 레벨에서도 프리페칭이 적용되어야 하는 점은 고려하지 않고, 2차원적 공간에서만 프리페칭을 수행하였다.

본 논문에서는 다단계 추상화 데이터로 구성된 웹 지리정보시스템에서 확대 질의와 축소 질의 시, 다른 레벨로의 이동을 고려하여 프리페칭이 적용되는 알고리즘을 제안하였다. 프리페칭 공간을 다단계로 확장시킨 알고리즘의 성능을 평가한 결과, 사용자의 응답시간이 1.8%~21.6% 빨라지는 성능 향상이 있었다. 따라서, 웹 지리정보시스템에 제안된 프리페칭 알고리즘과 이와 연동하는 캐쉬대체 알고리즘을 적용함으로써 사용자들에게 보다 빠른 응답시간으로 서비스할 수 있다.

* 가톨릭대학교 컴퓨터정보공학부 교수
** 인하대학교 전산정보원
*** 인하대학교 정보통신공학부 교수

ABSTRACT : The effective probability-based tile pre-fetching algorithm and the collaborative cache replacement algorithm are able to reduce the response time for user's requests by transferring tiles which will be used in advance and determining tiles which should be removed from the restrictive cache space of a client based on the future access probabilities in Web GISs(Geographical Information Systems). The Web GISs have multi-level abstracted data for the quick response time when zoom-in and zoom-out queries are requested. But, the previous pre-fetching algorithm is applied on only two-dimensional pre-fetching space, and doesn't consider expanded pre-fetching space for multi-level abstracted data in Web GISs.

In this thesis, a probability-based pre-fetching algorithm for multi-level abstracted in Web GISs was proposed. This algorithm expanded the previous two-dimensional pre-fetching space into three-dimensional one for pre-fetching tiles of the upper levels or lower levels. Moreover, we evaluated the effect of the proposed pre-fetching algorithm by using a simulation method. Through the experimental results, the response time for user requests was improved 1.8%~21.6% on the average. Consequently, in Web GISs with multi-level abstracted data, the proposed pre-fetching algorithm and the collaborative cache replacement algorithm can reduce the response time for user requests substantially.

Keywords : Web Geographic Information Systems, pre-fetching algorithm, collaborative cache replacement algorithm, multi-level abstracted data

1. 서 론

최근 인터넷의 대중화와 통신의 급격한 발달로 인해 기존의 지리정보시스템(Geographical Information System: GIS)을 확장한 웹 기반의 지리정보시스템이 발달하였다. 웹 기반 지리정보시스템(Web GIS)은 인터넷 사용자에게 다양하고 정확한 지리정보를 제공할 수 있으나, 사용자의 요구가 늘어나고 서비스 이용률이 높아짐에 따라 더욱 많은 양의 지리 데이터를 서버에서 클라이언트로 전송해야 할 필요성이 대두되고 있다(Annavaram과 Patel, 2001; Cha 등, 2000; Park등, 2000).

웹 지리정보시스템에서는 지리 데이터를 서버에서 클라이언트로 전송할 때, 전체 지리 데이터를 한꺼번에 전송하거나 또는 지리 데이터를 레이어 단위로 전송한다. 이 두 방식은 클라이언트가 레이어의 일부 영역에 포함되는 지리 데이터만을 요구하여도 서버로부터 전체 데이터를 전송 받거나 레이어 전체를 전송 받아야 하기 때문에, 클라이언트의 초기 시스템 지연 시간이 매우 크다. 이를 해결하기 위해 지도의 전체를 여러 개로 나눈 조

각, 즉 타일을 전송 단위로 사용하는 방법을 이용한다(Coetzee와 Bishop, 1998).

전송 단위가 작은 타일을 전송하는 방법은 서버가 클라이언트에 타일 단위로 전송하고, 클라이언트에서는 먼저 전송된 타일을 사용자에게 우선 보여줌으로써 초기 로딩 시간을 줄일 수 있다. 하지만 결국 지도를 보기 위해서는 요청한 타일들이 모두 도착하기를 기다려야 하므로 전체 응답시간은 전혀 줄지 않는다. 이러한 문제를 해결하기 위해서 제안된 것이 타일 프리페칭 알고리즘이다. 타일 프리페칭 알고리즘은 사용자가 앞으로 사용하게 될 타일들을 정확하게 예측하여 미리 클라이언트에 가져다 놓음으로써 사용자가 긴 네트워크 전송시간 동안 기다리는 대신에 빠른 응답으로 서비스를 받을 수 있게 한다(Tuah와 Venkatesh, 2001).

웹 지리정보시스템이 관리하는 지리 데이터는 단계별로 데이터의 복잡도를 다르게 한 다단계 추상화 데이터로 구성되어 있다. 사용자가 확대, 축소 질의를 요청한 경우에는 현재 지리 데이터를 포함하고 있는 단계가 아닌, 다른 단계의 지리 데이터가 필요하게 된다. 이와 같은 질의를 처리하기 위해 웹 지리정보시스템은 레벨링 기법을 사용하였고, 해당 레벨마다 데이터의 복잡도가 다른 다단계 추상화 데이터를 가진다. 다단계 추상화 데이터 구조를 가진 웹 지리정보시스템에서 일반적으로 사용자로부터 자주 요청되는 확대(zoom-in), 축소(zoom-out) 질의는 전혀 고려하지 않고, 점 질의, 이동 질의만을 고려한 2차원적인 타일 기반의 프리페칭 알고리즘은 현실성이 적다. 따라서, 기존 웹 지리정보시스템에서 점 질의, 이동 질의뿐만 아니라 확대, 축소 질의까지 고려한 확장된 프리페칭 알고리즘이 필요하다(Vakdi, 2000; Jiang과 Kleinrock, 1998).

본 논문에서는 사용자의 전체 응답시간을 줄이기 위해 Kang등 (2001)의 연구에서 제안한 확률기반모델로 표현된 사용자의 타일접근패턴을 이용한 2차원적인 프리페칭 알고리즘을 보완하여, 지도의 확대, 축소를 고려한 다단계 추상화 데이터의 프리페칭 알고리즘을 제안한다. 이 프리페칭 알고리즘은 레벨링 기법이 사용된 지리 데이터를 기반으로 하고, 프리페칭 단위는 타일을 기반으로 한다. 또한, 공간 데이터가 인접한 영역에 대한 질의가 많고 특정 영역에 질의가 집중된다는 특성을 이용하여, 사용자가 현재 검색중인 정보에서 공간적으로 멀리 떨어진 객체보다 가까이 존재하는 객체가 다음에 요청될 확률이 더 높다는 공간적 근접성을 기반으로 한다. 이와 연동하는 캐쉬 정책으로는 Kang등 (2001)의 연구에서 제안된 캐쉬대체 알고리즘을 사용하여 프리페칭의 효율성을 높였다.

본 논문의 구성은 다음과 같다. 2장에서는 타일링과 레벨링 기법을 설명하고 3장에서는 다단계 추상화 데이터의 프리페칭 알고리즘을 제안한다. 4장에서는 실험환경 및 알고리즘 성능평가를 하고, 마지막으로 5장에서는 결론과 향후 연구방향을 제시한다.

2. 타일링과 레벨링 기법

2.1 타일링

넓은 지역에 해당하는 자료를 컴퓨터에 입력하여 관리할 때 관리 목적상 작은 단위 면적으로 나누어 관리하는 것이 편리하므로 전체 대상지역을 작은 단위 면적으로 분할하여 관리할 때 각각의 작은 면적을 나타내는 지도를 타일이라 하며, 타일을 만드는 과정을 타일링(Tiling)이라 한다. 일반적으로 타일의 경계는 데이터베이스에서 사용이 편리하고 전반적인 시스템의 성능 향상에 기여할 수 있도록 설정되어야 한다. 타일을 기반으로 하는 데이터의 관리는 데이터 파일의 효율적인 검색기능을 제공함으로써 시스템의 제반성능을 향상시킬 수 있다. 반면 대상지역 전체를 추출해야 하는 경우에는 자료의 결합을 위한 과정에서 시간이 많이 소요되므로 처리 속도가 늦어진다는 단점이 있다.

일반적으로 지리 데이터는 행정경계, 지하철 노선, 건물, 도로 등과 같이 서로 다른 주제를 가지는 각각의 레이어로 분할되고, 각각의 레이어는 타일로 분할되어 관리된다. 타일로 분할된 각각의 레이어가 통합되면 새로운 지리 데이터가 생성된다(Kang등, 2001). 이렇게 생성되는 지리 데이터는 웹 지리정보시스템에서 타일 단위로 전송된다. 따라서, 프리페칭의 단위도 타일 기반이다.

2.2 레벨링

지도의 축척이 감소될 경우 시각적 복잡성이 나타나므로 지도에서는 공간을 체계적으로 단순화하여 나타내주어야 한다. 그리고 지리정보시스템에서는 레이어를 기반으로 한 지역에 대해 계속 확대를 하면서 객체의 상세한 정도를 결정한다. 전체적인 지형을 조회하기 위해 자세한 지리정보를 가진 한 개의 타일을 가져오는 것은 네트워크의 지연뿐만 아니라 사용자의 지형 판독가능성도 떨어뜨린다. 그러므로 다양한 용도에 맞는 최적화된 타일을 설계해야 하며 이를 위해 다수 개의 레벨을 두어 해당 레벨마다 데이터의 복잡도를 다르게 한다.

건물, 도로, 행정 구역 등 각각 다른 주제를 가진 타일 기반의 레이어들은 지리 데이터의 복잡도를 다르게 한 여러 개의 레벨로 존재하며, 각 레벨에서는 축척이 같은 서로 다른 레이어들이 하나의 통합된 레이어처럼 관리된다.

3. 다단계 추상화된 데이터의 프리페칭 알고리즘

프리페칭은 사용자가 원하는 정보를 예측하여 사용자가 질의 요청을 하기 전에 미리 서버에서 클라이언트로 필요한 정보를 전송하는 것이다. 일반적으로 서버와 클라이언트로 이루어진 시스템에서는 사용자가 정보를 요청할 때마다 클라이언트가 서버에 사용자의 질의를 전달하고, 서버는 요청한 데이터를 찾기 위해 서버가

가지고 있는 데이터베이스의 자료들을 검색한다. 그리고 검색 결과를 다시 클라이언트로 전송하여 사용자에게 보여준다. 클라이언트에서 사용자의 질의가 발생할 때마다 서버의 데이터베이스의 자료들을 검색하는 것은 사용자의 질의에 대한 응답시간이 길어지므로 비효율적이다. 따라서, 데이터의 처리시간을 줄이고 사용자의 응답시간을 줄이기 위한 방법으로 프리페칭 기법이 제안되었고, 웹 지리정보시스템에서도 이러한 프리페칭 기법을 적용하여 사용자에게 빠른 응답시간으로 서비스 할 수 있게 한다.

웹 지리정보시스템은 대용량의 지리 데이터를 빠르게 전송하기 위해 레이어를 작은 조각으로 분할한 타일을 전송한다. 또한 지도의 확대와 축소를 위해 레벨링 기법이 사용되어 축척이 다른 타일 기반의 다수 개의 레벨을 지원한다. 기존의 프리페칭 알고리즘은 타일 기반의 레이어에서 좌, 우, 위, 아래로의 이동 질의만을 고려한 2차원적인 알고리즘이므로 웹 지리정보시스템에서는 이를 확장시킨 프리페칭 알고리즘이 필요하다.

본 논문에서는 지도의 확대와 축소를 위해 다단계 추상화 데이터로 구성된 웹 지리정보시스템에 적용시킬 수 있는 프리페칭 알고리즘을 제안한다. 실질적으로 사용자가 원하는 지리정보를 전송하기 위해서는 점 질의나 이동 질의뿐만 아니라 확대(zoom-in), 축소(zoom-out) 질의까지 처리할 수 있어야 하므로, 제안하는 알고리즘의 프리페칭 공간은 현재 타일과 인접하여 있는 좌, 우, 위, 아래의 4개 타일들과 현재 타일을 확대시켜 나타낸 상세 레벨에 있는 타일, 현재 타일을 축소시켜 나타낸 개괄 레벨의 해당 타일까지로 확장시킨다.

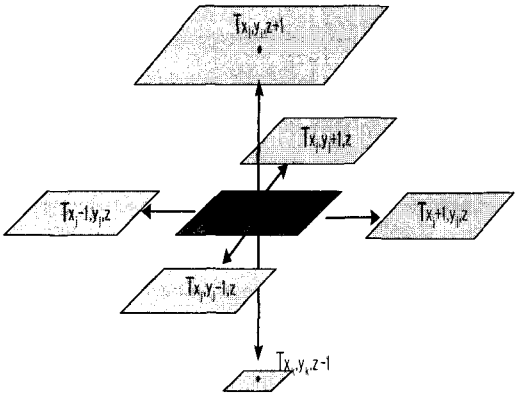
확대 질의와 축소 질의를 고려한 다단계 추상화 데이터에 프리페칭 알고리즘을 적용하기 위해서 웹 지리정보시스템은 최소 3개 이상 타일 기반의 레벨이 존재하여야 한다. 사용자의 질의 처리 결과로 클라이언트에 전송될 타일을 포함하는 레벨을 현재 레벨 또는 기준 레벨이라 하고, 현재 레벨을 중심으로 현재 레벨의 지리 데이터의 복잡도를 단순화시킨 레벨을 개괄 레벨, 현재 레벨의 지리 데이터를 상세히 나타낸 레벨을 상세 레벨이라 하며, 이 3개의 레벨이 프리페칭 알고리즘의 기본 단위가 된다.

사용자는 평균적으로 한번 접속하면 여러 개의 지리 데이터를 요청한다. 사용자 접근 패턴은 지리 데이터의 검색 및 프리페칭 시에 중요하다. 사용자의 지리 데이터 접근 패턴은 유동적이므로 일정기간동안 사용자의 타일 접근 패턴을 저장한 자료를 이용한다(Kang등, 2001).

프리페칭 알고리즘은 사용자의 타일접근패턴을 기반으로 다음 타일로 이동하게 될 확률을 계산하고, 계산된 이동확률 값에서 가장 큰 값을 가지는 타일을 프리페칭한다. 사용자는 현재 검색 중인 정보에서 공간적으로 멀리 떨어져 있는 타일보다는 가까이 존재하는 객체를 다음에 요청할 확률이 더 높다는 공간적 근접성을 기반으로 현재 사용중인 타일 다음에 사용할 타일을 결정한다. 프리페칭 알고리즘을 적용하기 위해 확장시킨 웹 지리정보시스템의 프리페칭 데이터베이스는 이러한 타일들 간의 이동확률을 저장하고

있다. 사용자에게 질의 처리 결과 타일을 전송한 후에 프리페칭 실행기가 프리페칭 데이터베이스를 이용하여 앞으로 사용자가 사용할 타일을 예측하고, 이동확률 값이 높은 타일을 프리페칭 한다.

본 논문에서 제안하는 다단계 추상화 데이터를 고려한 프리페칭 알고리즘은 이러한 타일 관계를 가진 각각의 레벨을 고려하여 프리페칭하는 알고리즘이다. 웹 지리정보시스템에서 확대 질의나 축소 질의를 처리하기 위해 최소 3개 이상의 레벨을 고려하여 프리페칭을 할 때, [그림 1]에서와 같이 2차원의 프리페칭 공간을 확장시킨 공간을 갖는다.



[그림 1] 제안한 알고리즘의 확장된 3차원 프리페칭 단위 공간

이러한 프리페칭 기본 공간을 사용하는 다단계 추상화 데이터의 프리페칭 알고리즘은 다음과 같다. 이 알고리즘은 사용자의 타일접근패턴을 기반으로 타일들의 이동확률 값을 계산하여 프리페칭을 수행하고 캐쉬 정책은 Kang등 (2001)의 연구에서 제안한 캐쉬대체 알고리즘을 사용한다. 캐쉬대체 알고리즘도 확장된 공간을 고려하여 다른 레벨로의 타일 이동확률 값도 계산하고 알고리즘을 수행한다.

[다단계 추상화 데이터의 프리페칭 알고리즘]

입력: 프리페칭 크기(pre-fetch_size), 캐쉬 타일목록(own_tile_list), 사용자의 타일 접근 패턴(local_access_pattern), 점 (a, b)를 포함하고 있는 타일 Tx,y,z를 중앙에 포함한 타일들

출력: 캐쉬대체를 위해 이동확률 값이 갱신된 캐쉬타일목록(own_tile_list), 프리페칭된 타일들과 결과 타일들

Step 1. 프리페칭 크기가 1보다 크거나 같은 경우에는 프리페칭을 수행하고, 프리페칭 크기가 1보다 작을 경우에는 프리페칭을 수행하지 않고 "NO_PRE-FETCH" 값을 리턴하고 끝낸다.

Step 2. 프리페칭 크기가 1이라면, 현재 타일로부터 프리페칭할 타일까지의 거리가 1인 위, 아래, 좌, 우의 타일에 대해서 정규화된 이동확률 값을 계산한다. 이 때, 확대 타일과 개괄 타일로의 이동확률 값은 정규화과정을 거치지 않고 사용자 타일접근패턴에 기반한 이동확률 값을 사용한다.

Step 3. 프리페칭 크기가 1보다 큰 경우에는 현재 타일에서 프리페칭할 타일까지의 거리가 프리페칭 크기 이하인 모든 타일에 대해서 이동확률 값을 계산한다.

Step 4. 계산된 이동확률 값들을 내림차

순으로 정렬하고 정렬된 이동확률 값들 중에서 이동확률 값이 높은 타일들을 프리페칭 크기만큼 구한다.

Step 5. 캐쉬 데이터를 갱신하기 위해 이전 질의에 의해 클라이언트에 캐쉬된 타일들은 프리페칭 타일목록에서 제거한다.

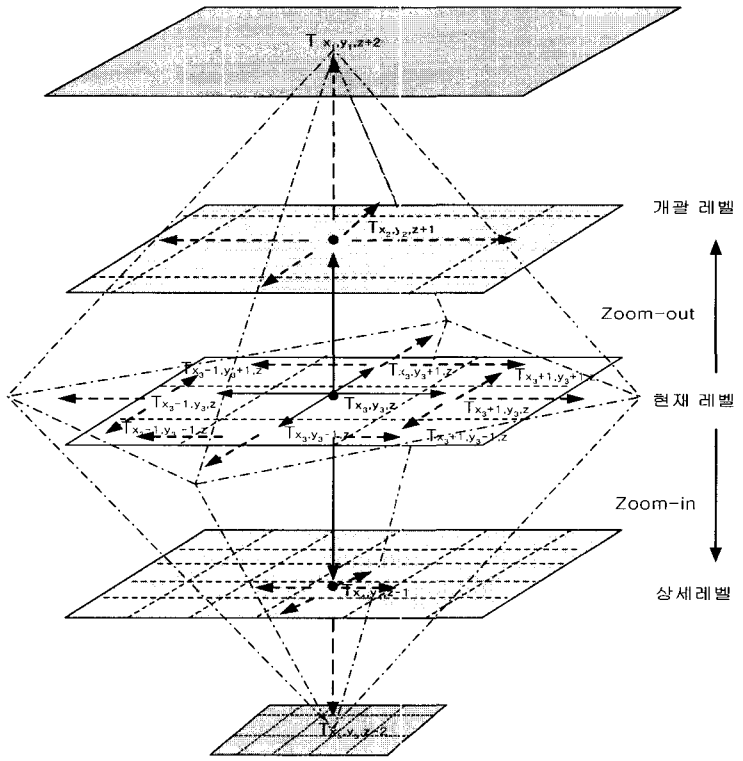
Step 6. 프리페칭될 타일의 목록과 타일들을 클라이언트에게 프리페칭의 결과로 돌려준다. 그리고 캐쉬 데이터베이스에 캐쉬된 타일 목록에 현재 결과 타일을 기준으로 하여 이동확률 값을 계산하고, 계산된 이동확률 값으로 캐쉬타일목록을 갱신하여 클라이언트에게 리턴한다. (프리페칭된 타일 목록과 이동확률 값이 갱신된 캐쉬타일목록을 리턴)

Step 7. 서버는 이동확률 값이 갱신된 캐쉬타일목록을 클라이언트의 캐쉬 관리자에게 보내준다.

서버에서 처리한 사용자의 질의 결과 타일이 클라이언트로 전송되어질 때, 서버는 사용자의 응답시간을 줄이기 위해 결과 타일이 전송되는 시간동안 미리 사용자가 사용하게 될 타일을 예측하여 클라이언트로 전송한다. 확률기반의 프리페칭 알고리즘을 수행하여 이동확률 값이 가장 높은 타일부터 내림차순으로 정렬한 후, 이 중에서 프리페칭 크기만큼 이동확률 값이 높은 타일들을 프리페칭한다. 프리페칭될 타일이 결정되면 서버는 프리페칭될 타일들의 목록과 타일, 그리고 이동확률 값이 갱신된 캐쉬 타일 목록을 클라이언트에게 보내줌으로써 사용자의 응답시간을 줄일 수 있는 효과를 가져온다.

다단계 추상화 데이터를 기반으로 하는 웹 지리정보시스템에서는 2차원적인 타일 프리페칭 공간과 상세 레벨의 타일과 개괄 레벨의 타일도 포함하는 확장된 3차원 프리페칭 공간을 갖는다. 웹 지리정보시스템에서 임의의 점 (a,b)에 대해 프리페칭을 수행할 때, 프리페칭 크기가 1이면 동일 레벨에서 현재 타일과 인접한 4개의 타일과 확대된 레벨에서 점 (a,b)를 포함하고 있는 타일, 축소된 레벨에서 점 (a,b)를 포함하고 있는 타일로 총 6개의 타일이 프리페칭 대상 타일이 된다.

프리페칭 크기가 2인 경우는 연속적인 확대 질의나 축소 질의를 요청한 경우에 3단계의 레벨이 필요로 하게 된다. 사용자에게 전송한 결과 타일 Txj,yj,z를 기준으로 가장 먼저 확대 질의나 축소 질의 처리가 요청된 경우는 현재 레벨에서 1단계 확대나 축소된 레벨로의 이동이 필요하다. 확대 질의가 발생한 경우라면, 결과 타일로부터 이동거리가 1인 점 (a,b)를 포함하는 상세 타일에서 다시 확대나 축소 질의를 고려한 프리페칭 수행이 필요하게 된다. 이 때 또 한번의 확대질의가 발생하면 상세 레벨에서 1단계 더 확대된 상세 레벨의 타일을 프리페칭할 대상 타일에 포함시키는 것이다. 프리페칭 크기가 2인 경우는 결과 타일 Txj,yj,z가 존재하는 레벨에서 이동거리가 1인 6개의 타일과 이동거리가 2인 18개의 타일을 포함하는 공간이 프리페칭 공간이다. 즉, 점 (a,b)를 포함하는 해당 레벨에서 12개의 프리페칭 대상 타일이 있고, 1단계 확대나 축소된 레벨에서 각각 5개의 타일, 즉 10개의 타일과 이동 거리가 2인 2단계 떨어진 레벨

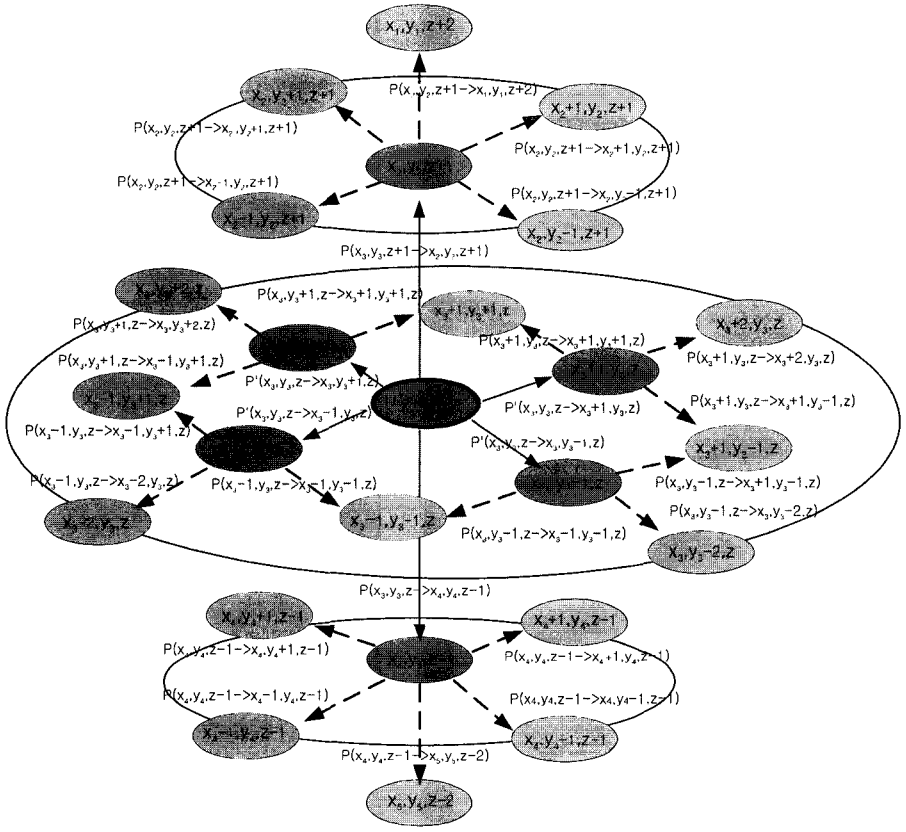


[그림 2] 프리페칭 크기가 2일때 확장된 3차원 프리페칭 공간

에서 각각 확대된 타일 1개, 축소된 타일 1개를 포함하여 총 24개의 타일을 프리페칭 대상으로 하고, 이 중 이동확률 값이 높은 2개의 타일만을 프리페칭한다. [그림 2]에서는 프리페칭 크기가 2인 경우의 3차원 프리페칭 공간을 나타낸 그림이다.

[그림 2]에서 보여주는 입체적인 프리페칭 공간에서 타일 $T_{x3,y3,z}$로부터 프리페칭할 타일을 선정하기 위해서는 타일 이동확률 값을 계산하고 이 중에서 가장 높은 이동확률 값을 가진 타일 순으로 프리페칭 크기만큼 타일을 프리페칭한다.

확대, 축소 질의를 고려한 프리페칭 알고리즘에서는 프리페칭 공간 안에서 가장 인접한 6개의 타일, $T_{x3+1,y3,z}$, $T_{x3,y3+1,z}$, $T_{x3-1,y3,z}$, $T_{x3,y3-1,z}$, $T_{x3,y3,z+1}$, $T_{x3,y3,z-1}$들은 한 번의 타일 이동으로 도달할 수 있다. 이는 프리페칭 크기가 1인 경우로 프리페칭 알고리즘을 수행했을 경우, 이 6개의 타일 중에 이동확률 값이 가장 높은 타일 1개만을 결정하여 클라이언트로 보내게 된다. 타일 $T_{x3,y3,z}$로부터 두 번의 타일 이동으로 도달할 수 있는 타일은 $T_{x3,y3+2,z}$, $T_{x3+1,y3+1,z}$, $T_{x3+2,y3,z}$, $T_{x3+1,y3-1,z}$, $T_{x3,y3-2,z}$, $T_{x3-1,y3-1,z}$, $T_{x3-2,y3,z}$, $T_{x3-1,y3+1,z}$, $T_{x2,y2+1,z+1}$, $T_{x2+1,y2,z+1}$, $T_{x2,y2-1,z+1}$, $T_{x2-1,y2,z+1}$, $T_{x4,y4+1,z-1}$, $T_{x4+1,y4,z-1}$, $T_{x4,y4-1,z-1}$, $T_{x4-1,y4,z-1}$, $T_{x1,y1,z+2}$, $T_{x5,y5,z-2}$인 18개의 타일들이다.



[그림 3] 프리페칭 크기가 2일때 프리페칭 공간안에서의 이동확률값

[그림 3]에서 이웃한 타일로의 화살표는 타일 이동을 의미하며, 화살표의 라벨은 이동확률 값을 의미한다. $P(x_3,y_3,z \rightarrow x_3,y_3+1,z)$은 타일 $T_{x3,y3,z}$에서 타일 $T_{x3,y3+1,z}$로의 이동확률 값을 나타낸다. 타일 $T_{x3,y3,z}$에서 좌, 우, 위, 아래 방향의 인접한 4개의 타일로의 이동확률 값은 정규화된 이동확률 값을 계산하고, 확대 레벨이나 축소 레벨로의 타일 이동확률 값은 정규화 과정을 거치지 않고 사용자 타일접근패턴에 기반한 이동확률 값을 사용한다. 정규화된 이동확률 값은 P'로 나타내고 일반적인 이동확률 값은 P로 나타낸다.

이동확률 값은 2차원적 프리페칭 알고리즘의 정규화된 계산식을 이용하며, 일반화된 식은 다음과 같다.

$$P(x,y \rightarrow x+n,y+m) = SUM(conditional\ probabilities\ of\ all\ paths\ from\ T_{x,y}\ to\ T_{x+n,y+m})$$

식(1)

위의 식을 다단계 추상화 데이터에도 프리페칭을 적용하기 위해서는 식의 $P(x,y \rightarrow x+n,y+m)$에서 $P(xm,yn,z \rightarrow xm+i,yn+j,z+k)$의 확장이 필요하다. 따라서, 다단계 추상화 데이터에서 프리페칭 적용시에 사용되는 일반식

은 다음과 같다.

$$P(x_m, y_n, z \rightarrow x_m+i, y_n+j, z+k) = SUM(conditional\ probabilities\ of\ all\ paths\ from\ Tx_m, y_n, z\ to\ Tx_m+i, y_n+j, z+k)$$

식(2)

타일 $T_{xm,yn,z}$에서의 점 (a,b)라는 위치를 중심으로 위, 아래, 좌, 우로 움직이는 이동질의 처리시 타일들의 이동확률 값을 구하기 위해서는 조건부(conditional) 이동확률 값을 계산하나, 상세레벨이나 개괄레벨로의 이동시에는 정규화한 이동확률 값을 사용하지 않는다. 프리페칭 크기가 2인 경우에 타일 $T_{xm,yn,z}$로부터 거리가 2인 타일의 이동확률 값을 계산하려면, 먼저 거리가 1인 타일로의 이동확률 값을 구하고 이동확률 값을 구한 타일에 대해 또 한번 이동거리가 1인 타일로의 이동확률 값을 곱한 값들을 합하여 이동확률 값을 계산한다.

계산한 이동확률 값들을 내림차순으로 정렬하고, 그 중에서 이동확률 값이 높은 순으로 프리페칭할 타일을 2개 결정하여 클라이언트에 전송한다. 이와 연동하는 캐쉬 정책으로 Kang등 (2001)의 연구에서 제안된 캐쉬대체 알고리즘을 사용하여 프리페칭의 효율성을 높여 사용자의 응답시간을 줄인다.

4. 확장된 프리페칭 알고리즘의 성능 평가

4.1 성능 평가 환경

다단계 추상화 데이터의 프리페칭 알고리즘을 수행하기 위해서 웹 지리정보 시스템은 지리 데이터의 복잡도가 다른 여러 개의 레벨들로 구성되어야 하므로 실험에서도 이를 반영하여 7개의 레벨을 두어 실험하였다. 알고리즘의 성능 평가를 위해서 사용된 시스템은 주기억장치 256MB, Pentium Ⅳ 1GHz이다. 실험에 사용된 운영체계는 Windows 2000 Professional, 시뮬레이션 도구는 C++ SIM 5.0, 개발도구는 MS Visual C++ 6.0을 이용하였다[9].

제안한 확률기반 프리페칭 알고리즘의 효율성을 입증하기 위해 프리페칭 알고리즘과 캐쉬대체 알고리즘, 이 두 가지 알고리즘이 적용되지 않은 기존 웹 지리정보시스템에서 실험하여 사용자 응답시간을 비교하였다. 확률 기반의 프리페칭 알고리즘과 캐쉬대체 알고리즘을 평가하기 위해 설정한 변수값들은 다음 <표 1>과 같다. 실험에서 평균 디스크 액세스 시간은 25ms로 가정하였다.

<표 1> 실험을 위해 설정한 변수값

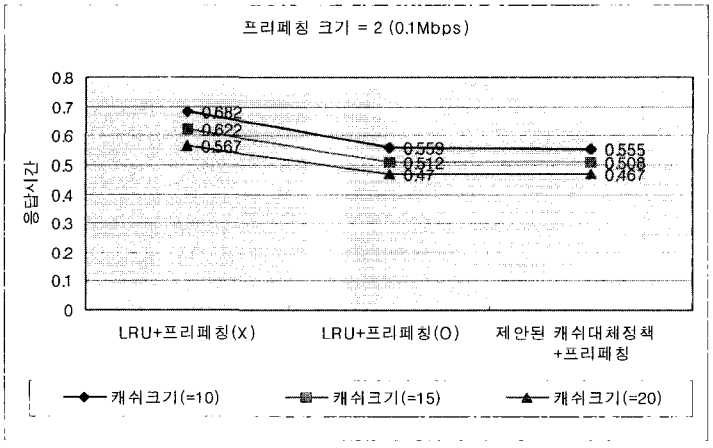
알고리즘 수행 기본 항목	설 정 값
레벨 1에서의 지리데이터	3 × 3 타일 구조
레벨 2에서의 지리데이터	5 × 5 타일 구조
레벨 3에서의 지리데이터	7 × 7 타일 구조
레벨 4에서의 지리데이터	9 × 9 타일 구조
레벨 5에서의 지리데이터	10 × 10 타일 구조
레벨 6에서의 지리데이터	11 × 11 타일 구조
레벨 7에서의 지리데이터	12 × 12 타일 구조
레벨의 타일 크기	9kbytes
캐쉬 정책 사용시의 캐쉬 데이터베이스의 크기	지리정보 데이터베이스의 10%, 15%, 20%
질의 발생시간	1초
네트워크 전송속도	0.1, 0.5, 1, 3, 5, 8, 10 Mbps

제안된 프리페칭 알고리즘에 대한 성능 평가는 3가지의 경우를 고려하여 수행하였다. [경우 1]은 웹 지리정보시스템에서 지리정보 검색 시에 프리페칭을 하지 않고 캐쉬대체 알고리즘으로는 기존의 LRU를 사용하는 경우이다. [경우 2]는 본 논문에서 제안한 타일 프리페칭 알고리즘과 캐쉬대체 알고리즘으로는 LRU를 적용한 경우로서 두 가지 알고리즘이 적용되지 않은 경우와 프리페칭만을 수행한 경우의 비교를 위한 것으로 프리페칭 알고리즘의 성능을 평가한다. [경우 3]은 본 논문에서 제안한 타일 프리페칭과 캐쉬대체 알고리즘 모두 적용한 경우로 [경우 2]와 비교할 때 캐쉬대체 알고리즘의 성능을 확인할 수 있다.

알고리즘 성능 평가에 사용된 사용자 질의의 종류는 점 질의와 이동 질의, 확대 질의, 축소 질의이다. 이는 기존에 제안된 2차원적 타일 프리페칭 알고리즘에서 적용하지 못한, 확대 질의와 축소 질의 처리를 포함한 것으로 현실에 가까운 웹 지리정보 시스템을 구성한 것이다. 성능 평가는 사용자가 임의의 점 (a,b)를 찾는 점 질의를 요청한 후에 이동 질의, 확대 질의, 축소 질의가 요청된다고 가정하였다. 10~30사이의 임의의 수만큼 이동 질의와 확대 질의, 축소 질의를 한 후에 다시 임의로 점 질의를 하고 다시 10~30사이의 임의의 수만큼 이동 질의, 확대 질의, 축소 질의를 하는 것을 반복하여 100,000번의 질의를 수행하였다. 또한 사용자의 질의는 전체 사용자의 타일접근패턴을 기반으로 하여 임의로 발생하도록 하였다.

4.2 알고리즘의 성능 평가 결과

Kang등 (2001)의 연구에서 제안한 2차원적 타일 프리페칭 알고리즘을 3차원으로 확장시킨 타일 프리페칭 알고리즘과 캐쉬대체 알고리즘의 성능을 평가한 3가지의 경우에 대해 결과를 보여준다. 실험환경에서 설정한 0.1, 0.5, 1, 3, 5, 8, 10Mbps의 네트워크 전송속도를 모두 적



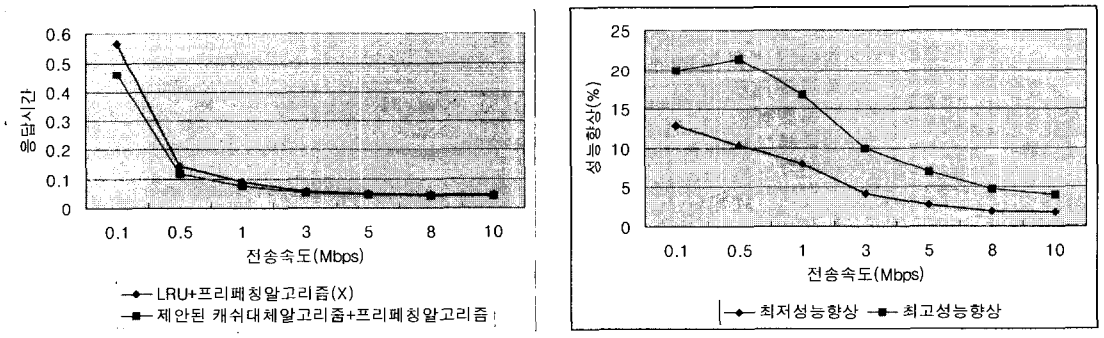
[그림 4] 세 가지 경우에 대한 응답시간

용한 결과 데이터 중에서 [그림 4]는 전송속도가 0.1Mbps이고 프리페칭 크기가 2일 때 세 가지 경우에 대한 응답시간의 변화를 나타나는 실험 결과이다.

X축은 3가지의 경우인 프리페칭 알고리즘과 캐쉬대체알고리즘을 모두 적용하지 않은 경우, 프리페칭 알고리즘과 LRU를 적용한 경우, 두 가지 알고리즘을 모두 적용한 경우를 나타내고 Y축은 사용자 응답시간을 나타낸다. 각각 3가지의 경우에 따라 클라이언트의 캐쉬의 크기가 10%, 15%, 20%일 때 변화하는 응답시간을 나타낸다. 서버에서 프리페칭 알고리즘을 수행하고 클라이언트에서 캐쉬대체 알고리즘이 수행된 경우가 프리페칭 알고리즘과 캐쉬대체 알고리즘을 모두 사용하지 않은 경우보다 응답시간이 12.9%~19.9%정도 향상되었다. 프리페칭 알고리즘은 수행되고 캐쉬대체 정책으로 LRU가 사용된 경우와 제안된 캐쉬대체 알고리즘을 사용한 경우를 비교하면 약 0.4%~0.8% 정도의 차이로 캐쉬대체 알고리즘을 사용한 경우가 사용자의 응답시간이 향상되었음을 보여주었다.

위의 실험 결과를 통해 0.1, 0.5, 1.0, 3, 5, 8, 10Mbps의 네트워크 전송속도일 때 제안된 두 가지 알고리즘을 사용했을 경우가 사용하지 않을 경우보다 1.8%~21.6%정도 빠른 응답시간을 보였다. 서버에서 프리페칭 알고리즘을 수행한 경우와 수행하지 않은 경우의 응답시간이 크게 차이가 나는 것을 알 수 있으므로 웹 지리정보시스템에서 프리페칭 알고리즘이 적용되었을 경우에 보다 빠른 응답시간을 서비스한다는 것을 입증한다. 그러나 캐쉬대체 알고리즘의 성능은 전송속도가 빨라질수록 캐쉬대체 알고리즘을 적용한 경우와 그렇지 않은 경우의 응답시간에는 약간의 성능 향상만을 볼 수 있었다. 제안된 확률 기반의 프리페칭 알고리즘과 캐쉬대체 알고리즘을 사용했을 경우 사용자의 응답시간은 캐쉬 데이터베이스의 크기와 프리페칭 크기에 따라 영향을 받는다.

[그림 5]는 전송속도에 따른 응답시간과 응답시간의 성능향상을 나타낸 그림이다. [그림 5]의 (a)는 프리페칭 크기가 3이고, 캐쉬 데이터베이스의 크기가 서버 측



(a) 전송속도에 따른 응답시간 (b) 전송속도에 따른 응답시간의 성능향상

[그림 5] 전송속도에 따른 응답시간의 변화와 성능향상

의 지리정보 데이터베이스 크기의 20%일 때, 전송속도에 따른 실험 결과로 프리페칭 알고리즘을 적용한 경우와 적용하지 않은 경우의 응답시간의 차이를 보여준다. [그림 5]의 (b)는 프리페칭 알고리즘과 캐쉬대체 알고리즘을 모두 적용했을 때의 전송속도에 따른 응답시간의 최저성능향상과 최고성능향상을 보여준다.

전송속도가 0.5Mbps이고 프리페칭 알고리즘과 캐쉬대체 알고리즘을 적용했을 때, 최고의 응답시간 성능향상을 보여주며 전송속도가 빠를수록 알고리즘의 성능차이가 거의 없다. 이러한 이유는 데이터 전송속도가 빨라질수록 클라이언트 디스크 액세스 시간과 거의 비슷하기 때문이다.

다단계 추상화 데이터를 가지는 웹 지리정보시스템에서 서버 측에서 확장된 알고리즘을 수행하고 클라이언트 측에서 캐쉬대체 알고리즘을 수행할 시, 실험결과를 통해 제안된 프리페칭 알고리즘과 캐쉬대체 알고리즘을 모두 적용했을 경우가 적용되지 않은 경우보다 빠른 응답시간을 가지고, 프리페칭 알고리즘을 수행시에는 프리페칭 크기가 클수록, 캐쉬대체 알고리즘을 수행시에는 클라이언트 측의 캐쉬 데이터베이스의 크기가 클수록 빠른 응답을 보인다는 것을 증명하였다.

5. 결론 및 향후 연구방향

웹 기반 지리정보시스템은 사용자에게 다양하고 정확한 지리정보를 제공하고, 늘어나는 사용자의 요구를 만족시키기 위해서는 더욱 많은 양의 지리 데이터를 서버에서 클라이언트로 전송해야 할 필요성이 대두되고 있다. 기존의 웹 지리정보시스템에서 사용자가 지리정보를 검색하고자 할 때 지도의 초기 로딩시간이 오래 걸리고 다양한 질의 처리에 대한 응답시간이 느린 단점을 가지고 있어, 이를 해결하기 위한 연구들이 많이 진행되고 있다.

사용자 질의에 대해 늦은 응답시간을 줄이는 것에 초점을 맞춰 제안된 알고리즘으로 프리페칭 기법이 있다. 이 프리페칭 알고리즘은 서버로부터 사용자가 가까운 미래에 사용하게 될 타일들을 미리 예측하여 클라이언트로 가져다 놓음으로써 사용자에게 빠른 응답시간으로 서비스를 받게 한다. 그러나 기존에 제안된 2차원적인 프리페칭 알고리즘은 단지 점 질의, 이동 질의만을 고려한 것으로서 일반적으로 많이 사용하는 확대 질의와 축소 질의를 처리하지 못하므로 현실성이 적다. 따라서, 확대 질의와 축소 질의도 고려한 확장된 프리페칭이 필요하다. 웹 지리정보시스템에서 확대 질의와 축소 질의를 빠르게 처리하기 위해서는 레벨링 기법을 사용한 다단계 추상화 데이터가 존재하고, 이 다단계 추상화 데이터를 고려하여 프리페칭이 적용되어야 한다.

본 논문에서는 다단계 추상화 데이터를 지원하는 웹 지리정보시스템에서 2차원적인 프리페칭 알고리즘을 확대 질의, 축소 질의를 모두 고려하는 확장된 3차원 프리페칭 알고리즘을 제안하였고, 시뮬레이션을 통해 제안한 알고리즘의 효율성을 입증하였다. 본 논문에서는 임의의 점 질의 발생 후 10~30번의 연속되는 이동 질의,

확대 질의, 축소 질의에 대해 실험을 하였고, 제안된 프리페칭 알고리즘을 사용하였을 때 사용자의 응답시간의 1.8%~21.6%정도 단축되는 성능향상이 있었다.

효율적인 다단계 추상화 데이터의 확률기반 프리페칭 알고리즘은 앞으로 웹 지리정보시스템의 검색시 보다 빠른 사용자의 응답시간을 서비스함으로써 웹 지리정보시스템의 발전에 많은 영향을 줄 것이다. 본 논문에서는 웹 지리정보시스템에서 사용하는 사용자 질의 중 영역 질의와 개체 추출 질의를 고려하지 않았으나 향후에는 이러한 질의들도 고려하는 프리페칭 알고리즘이 필요하다.

참고문헌

Annavaram, M. and Patel, J. M. 2001. "Call Graph Prefetching for Database Applications," Proc. of 7th Int. Symp. on High-Performance Computer Architecture, pp.281-290.

Cha, S. K. et al., 2000, "Efficient Web-based Access to Multiple Geographic Databases through Automatically Generated Wrappers," Proc. of 1st Int. Conf. on Web Information Systems Engineering, pp.34 -41.

Coetzee, S. and Bishop, J., 1998, "A New Way to Query GISs on the Web," IEEE Software.

Jiang, Z. and Kleinrock, L., 1998, "Web Prefetching in a Mobile Environment," IEEE Personal Communications, Vol.5, pp.25-34.

Kang, Y. K. et al., 2001, "Probability-Based Tile Pre-fetching and Cache Replacement Algorithms for Web Geographical Information Systems," Lecture Notes in Computer Science, Springer.

Little, M. and McCue, D., 1993, "Construction and Use of a Simulation Package in C++," Tech. Rep.437, Dept. of Comp. Sci., U. of Newcastle.

Park, G. H., Han, T. D. and Kim, S. D., 2000, "Analyzing Instruction Prefetching Techniques via a Cache Performance Model: Effectiveness and Limitations," Proc. of IEEE Int. Conf. on Performance Computing and Communications, pp.501-508.

Tuah, N. J. and Venkatesh, M., 2001, "Effect of Speculative Prefetching on Network Load in Distributed Systems," Proc. of 15th Int. Symp. on Parallel and Distributed Processing.

Vakali, A., 2000, "Data Block Prefetching and Caching in a Hierarchical Storage Model," Aristotelian Univ Salonika, SOURCE Information Sciences, Vol.128 No.1-2, pp.19-41.