GIS와 Web을 통합한 모바일 지역정보 검색 시스템*

강용진[○†] Yutaka Kidawara[‡] 권용진[†‡] Katusmi Tanaka[‡]
한국항공대학교 정보통신공학과[†]
Interactive Comm. Media and Contents Group, Keihanna Human Info-Comm. Research Center[‡]
{jupiterk[○], yjkwon}@tikwon.hangkong.ac.kr[†]
kidawara@nict.go.jp[‡] tanaka@dl.kuis.kyoto-u.ac.jp[‡]

Mobile Regional Information Retrieval System of Convergence of GIS and Web

Yongjin Kang[○†] Yutaka Kidawara[‡] Yong-Jin Kwon[†‡] Katusmi Tanaka[‡]
Dept. of Info. & Telecom. Eng., Hankuk Aviation University[†]
Interactive Comm. Media and Contents Group, Keihanna Human Info-Comm. Research Center[‡]

요 약

인터넷 기술의 발전과 웹 사용자의 증가로 웹에는 매우 많은 정보들이 축적되고 있으며, 여기에는 또한 매우 많은 지역정보가 존재한다. 일반적으로 웹 정보라고 하는 것은 웹의 컨텐츠, 사용이력, 구조를 중심으로 한 정보이며, 이러한 정보를 하나의 중요한 공간 정보원으로 취급하여 웹으로부터 특정 지역에 관한 컨텐츠를 수집하고 이를 분석하여 사용자의 공간 정보원으로서 사용할 수 있다. 또한 인터넷, 무선통신, 지리정보시스템(GIS), 위치기반서비스(LBS) 등의 기술들이 모바일기기에 복합적으로 접목되면서 모바일기기에서의 새로운 형태의 정보에 대한 응용이 요구되고 있다. 본 논문에서는 이러한 웹 환경을 활용하고, 모바일 환경을 고려하여 GIS와 Web을 통합한 모바일 지역정보 검색 시스템을 제안한다.

1. 서 론

인터넷 기술의 발전과 웹 사용자의 증가로 인터넷에는 매우 많은 정보들이 축적되었고, 또한 축적되고 있다. 그러나 이러한 웹 정보들은 많고 다양하다는 장점에 반해 형식적이거나 정량적으로 정제되어 있지 않다는 단점이 있어 이를 보완하기 위한 연구가 필요하며, 해당 연구들이 정보검색 분야에서 활발히 진행 중이다. 또한 인터넷, 무선통신, 지리정보시스템(GIS), 위치기반서비스(LBS) 등의 기술들이 모바일기기에 복합적으로 접목되면서 새로운 형태의 정보에 대한 응용이 요구되고 있다. 특히 무선 인터넷 기술과 컴퓨터 성능에 버금가는 PDA 및 Handheld PC와 같은 모바일기기의 등장은 사용자로 하여금 언제 어디서나 원하는 정보를 쉽게 얻을 수 있는 환경을 제공해주었으며, 이를 응용한 다양한 서비스들 또한 연구, 개발되고 있다. GIS(Geographic Information System) 분야에서도 이와 같은 모바일 기술이 응용되어 Mobile GIS라는 새로운 분야를 등장시켰으며, 인터넷 기술과 융합하여 Web GIS라는 분야를 만들어냈다. Mobile GIS란 GIS 기술에 모바일 기술을 결합한 형태로 별도의 장치나 시간, 공간상의 제약이 없이 사용자가 필요로 하는 지리정보를 실시간으로 사용자에게 제공하는 시스템을 의미하며, 사용자 중심의 위치기반기술의 형태로 발전하고 있다. 또한 Web GIS 역시 Web 환경과 기술의 발달과 더불어 발전하고 있으며, Web의 방대한 데이터를 보다 효율적으로 이용하기 위해 웹마이닝(Web-Mining) 기술이 접목되고 있다.

특히 다음의 어느 관광객의 예와 같은 경우에 Web GIS에 효과적인 웹마이닝 기술의 응용이 요구된다. 만약 어느 관광객이 경주를 관광하려고 하는 경우, 인터넷 상에서 필요한 정보를 얻고자 한다면, 꽤나 많은 작업을 필요로 하게 된다. 우선 경주를 찾아가기 위해 경주로 가는 길이나 교통정보 등을 검색해야 하고, 경주에서 가볼만한 관광명소나 문화유산이 어디에 위치하는지를 조사해야 한다. 또한 각 관광명소마다 정확한 위치와 입장료 및 부대시설 등을 추가로 검색해야만 한다. 예를 들어 가볼만한 문화유산으로 "석굴암"이 검색됐을 때, 관광객은 "석굴암"에 가보기 위하여 인터넷에서 "석굴암"의 위치를 알 수 있는 지도정보를 검색하게 된다. 지도 검색과정에서 관광객은 "석굴암" 부근의 다른 문화유산인 "불국사"를 발견하게 되고, "불국사"에 흥미를 갖게 되어 "불국사"를 인터넷에서 검색하게 된다. 이와 같은 연관된 지역정보를 얻기 위해서 사용자는 추가적으로 검색을 시도해야만 한다. 위와 같은 경우 일련의 연관된 웹 검색들을 좀 더 효율적으로 할 수 있는 지리정보 검색 시스템이 필요하며, 본 논문에서는 Web 정보를 효과적으로 분석하고 이를 GIS에 통합하여, 연관된 지리정보 검색을 모바일기기에서 간편하고 효율적으로 수행할 수 있는 시스템을 제안한다.

본 논문의 2장 에서는 시스템의 전반적인 구조와 동작을 설명하고, 3장에서는 어떻게 웹 정보를 활용하는지를, 4장에서는 분류된 세 가지 공간 사이의 상호 연관성을 설명하고, 5장에서는 검색된 정보와 계산된 연관성을 제한적인 모바일 환경에서 어떻게 효율적으로 나타낼 수 있는지를 설명한다. 끝으로 6장에서는 결론을 맺고 향후 연구과제에 대해 언급한다.

2. System Overview

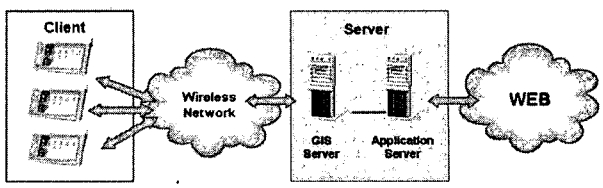

그림1. System Overview

* 본 논문은 산업자원부 한국산업기술평가원 지정 한국항공대학교 부설 인터넷정보검색 연구센터의 지원에 의함.

그림1 은 본 논문에서 제안한 지역정보 검색 시스템의 전반적인 동작 구조를 나타낸다. 본 시스템은 모바일기기들로 이루어진 클라이언트(Client) 부분과 서버(Server) 부분으로 나뉜다.

클라이언트의 각 모바일기기들은 본 논문에서 제안한 모바일 환경에 적합한 검색 소프트웨어를 탑재하고 있으며, 무선 네트워크 환경에서 서버와 상호 통신을 한다.

서버 부분은 GIS 서버와 응용(Application) 서버로 구분되며, GIS 서버는 지도정보 등의 지리정보에 대한 처리를 하고, 응용 서버는 웹으로부터의 정보 추출 및 상호 연관성 계산 등의 전처리 작업을 수행한다.

응용 서버는 지도정보로부터 추출한 지리정보로부터 웹 정보를 필터링하여 수집하고 분석한 후, 이를 데이터베이스화해서 저장한 뒤, 검색 요청이 발생하면 분석된 웹 기반 지역정보를 클라이언트에게 제공한다. 이 때, GIS 서버 역시 검색 요청이 들어온 지역과 그 주변의 지도정보를 클라이언트로 보내주게 된다. 두 가지 정보를 받은 클라이언트는 탑재된 소프트웨어를 통해 지도정보와 분석된 웹 정보를 사용자에게 보여주게 된다. 분석된 웹 정보에는 사용자가 검색요청을 한 질의어와 관련 있는 지명단어들과 비지명단어들 그리고 웹 페이지들의 주소정보가 있다. 이들 정보를 웹으로부터 수집하고 분석하는 과정은 이어지는 3장과 4장에서 설명한다.

3. 웹으로부터 지역관련 정보를 추출

인터넷의 발달로 인해 웹에는 매우 많고 다양한 정보들이 축적되고 있으며, 여기에는 또한 매우 많은 지역정보가 존재한다. 이러한 웹 정보로부터 지역관련 정보들을 찾아낼 수 있으며, 이러한 지역관련 정보들만을 추출해 낼 수 있다. 그러나 인터넷에는 방대한 정보들이 무질서하게 널려 있게 때문에 원하는 지역정보만을 추출하고 이를 활용하기 위해서는 정제과정을 거쳐야 한다. 이처럼 무질서한 웹 정보를 정제하기 위해 Geoword라는 지역관련 명사들을 선정하고 이들 집합을 필터로써 활용하여 지역관련 웹 정보만을 추출할 수 있다. 그러나 필터로써 활용할 Geoword들의 집합을 얻는 것은 간단한 작업은 아니며 매우 많은 작업량을 필요로 한다. 이처럼 Geoword 집합을 생성할 때 발생하는 과도한 부하는 전체 시스템의 성능을 현저히 떨어뜨리므로 이를 해결하기위해 획기적인 Geoword 집합 생성 알고리즘이나 방법이 필요하다.

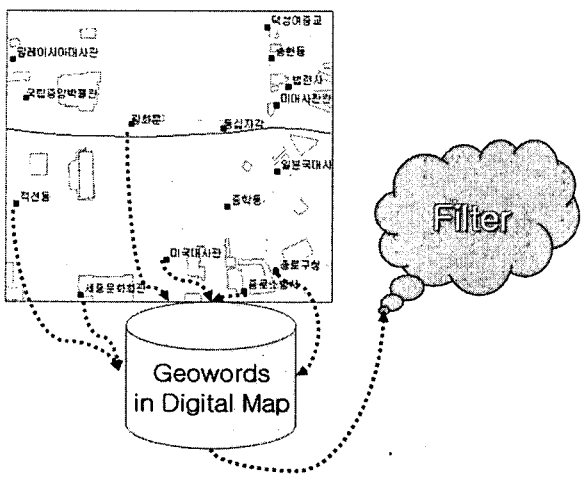

그림2. Geoword의 추출과 활용

건설교통부 산하 국토지리정보원에서 제작 및 배포하는 수치지도(Digital Map)에는 산, 강, 도로, 건물 등의 지리정보 이외에도 행정구역명이나 고유지명 등을 나타내는 텍스트정보가 존재한다. 특정지역의 수치지도로부터 이들 텍스트 정보를 추출하면 추출된 텍스트들은 해당 지역의 지명 단어들을 모아놓은 집합이 되며, 본 시스템에서는 이를 Geoword로 활용하였다. 이처럼 정형화되어 있는 수치지도라는 데이터를 활용함으로써 Geoword 집합 생성에 소요되는 부하를 크게 줄였으며, 더불어 인간의 사고 없이 기계적으로 Geoword의 생성이 가능하도록 하였다. 위의 그림2 는 수치지도로부터 Geoword를 추출하고 이를 필터로 활용함을 보여준다. 그림2 처럼 수치지도는 지명명사들의 텍스트정보를 포함하고 있기 때문에 특정 지역의 수치지도만 주어지면 수치지도 데이터로부터 Geoword 집합을 쉽게 얻을 수 있다.

4. 지역적인 특징, 비지역적인 특징 그리고 웹 정보 사이의 상호 연관성

Geoword를 필터로 하여 웹으로부터 관련 지역정보를 수집하였다면, Geoword 집합과 수집된 웹 정보로부터 지역적인 특징을 나타내는 Geoword 공간, 비지역적인 특징을 나타내는 Non-Geoword 공간, 웹의 본래 특징을 갖는 웹 공간의 세 종류의 공간을 이끌어 낼 수 있다. Geoword 공간은 Geoword 집합으로 이루어진 공간이고, Non-Geoword 공간은 수집한 웹들이 가지고 있는 단어 중 Geoword 집합을 제외한 나머지 단어들의 집합으로 이루어진 공간이다. 웹 공간은 웹의 본래 특징을 그대로 유지하고 있는 원래의 페이지를 나타내는 URL주소들의 집합으로 이루어진 공간이다. 이들 세 공간에 속한 원소들 사이에는 상호 연관성이 존재하며, 이를 특정한 알고리즘을 사용하여 계산해 낼 수 있다. 또한 이러한 계산과정을 통하여 지역성 및 상호 연관성을 고려한 지역정보 검색이 가능한 시스템이 구축된다.[1]

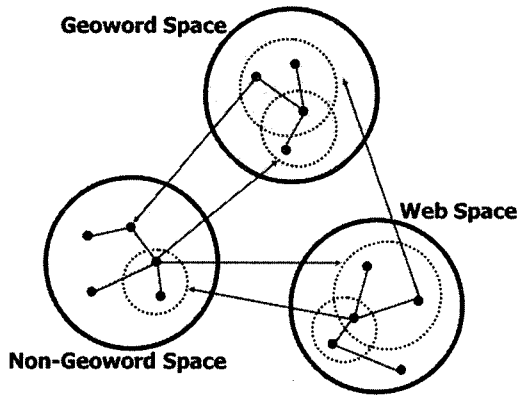

그림3. 세 종류의 공간간의 상관관계

그림3 에서는 Geoword 집합과 이를 활용해서 얻은 웹 기반 지역정보들로부터 생성해 낼 수 있는 세 종류의 공간과 이들 사이의 상관관계를 도식화한 것이다. 각 공간에 속한 원소들은 동일한 공간 내부의 특정 원소들 혹은 다른 공간의 특정 원소들과 서로 연관성이 있다. 예를 들어 경기도 고양시에 있는 "서오릉"이라는 지명과 "서오릉"을 나타내는 다섯 능의 이름인 "창릉", "홍릉", "경릉", "익릉", "명릉"은 서로 연관성이 있으

며 이들은 모두 Geoword 공간이라는 동일한 공간에 속한다. 또한 Geoword 공간의 "서오릉"과 Web 공간에 있는 "서오릉"에 대한 설명이 들어있는 웹페이지의 URL 주소도 서로 연관성이 있다. 이처럼 각 공간의 원소들 사이에는 서로 상관관계가 존재하며, 이들을 특정 알고리즘을 사용하여 정량적으로 판단하고 계산하여 순서화 시킬 수 있다. 본 논문에서 제안한 시스템에서는 이들의 연관성을 파악하기 위해 연관규칙과 연관클러스터를 사용하였으며, 이들의 계산의 계산량을 줄이기 위해 최적화된 희소행렬 계산방식을 사용하였으며[2,3], 이 계산을 통해 상호간의 연관성을 파악하고, 파악된 연관도를 순서화 했다. 이들을 순서화함으로써 특정 단어와 관련성이 있는 원소들 중 가장 중요한 원소들만을 선택할 수 있도록 했다.

5. 모바일 환경에 적합한 새로운 검색 방법 및 사용자 인터페이스

모바일기기들은 그 특성상 제한된 사용자 인터페이스를 가지고 있다. 예를 들어 작은 해상도의 디스플레이 장치라든지 실제 키보드나 마우스가 아닌 터치스크린과 스타일러스 펜을 활용한 입력장치가 그 대표적인 예이다. 본 논문에서는 이러한 모바일기기들의 제약사항을 극복하기 위해 모바일 환경에 적합한 검색 인터페이스를 제안한다.

우선 다양한 검색결과에 대한 정보를 모두 표현하기 위해 Java의 CardLayout[5] 형태의 UI(User Interface)를 구현하고, 디스플레이 화면을 좌우로 나누어 각 레이어(Layer)들 간의 연관성을 같이 확인할 수 있도록 했다. 또한 입력장치의 제약을 극복하기 위해 터치스크린과 스타일러스 펜을 적극 활용하여 Keyword리스트에 표시되는 연관단어를 클릭하거나 지도에 표시된 지명 명사를 클릭하면 해당 키워드나 해당 지명을 쉽게 검색할 수 있도록 하였다.[4]

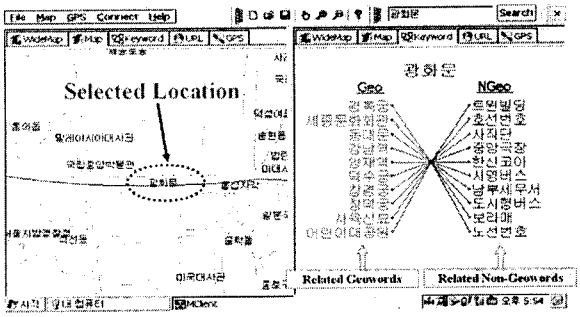

그림4. 모바일 환경을 고려한 시스템의 검색 UI

그림4 는 본 논문에서 제안한 모바일 환경을 고려한 검색 시스템의 검색 UI를 보여준다. 위의 예에서는 "광화문"이 검색되었으며, 화면 왼편에는 "광화문"과 그 주변의 지도정보가 표시되고, 오른편에는 "광화문"과 관련된 지명단어와 비지명단어가 관련성 상위 10개씩 표시되었음을 볼 수 있다. 이처럼 개선된 UI를 통해 모바일기기의 제한적인 사용자 환경을 최대한 극복하였다.

6. 결론 및 향후 연구

정점 비대해지는 Web이라는 정보공간에 지리정보시스템, 이동통신 등의 여러 기술들이 복합적으로 접목되어 새로운 형태의 데이터 응용이 요구되는 Mobile 환경 속에서 관련 기술들은 나날이 발전하며 서로 다른 기술, 서로 다른 데이터들 사이의 컨버젼스(Convergence)가 이루어지고 있다. 본 논문에서 제안한 "GIS와 Web을 통합한 모바일 지역정보 검색 시스템" 역시 Web과 GIS 와의 컨버젼스를 추구하는 시스템으로서, 방대한 양의 정보를 가지고 있는 Web을 정보원으로 활용하고, 수치지도라는 웹 정보와는 다른 형태의 정형화된 데이터를 이용하여 웹 정보를 정제하고 활용했다는 점에서 서로 다른 종류의 데이터, 다른 종류의 공간사이의 컨버젼스를 꽤한 시스템이라고 할 수 있다. 또한 수치지도로부터 추출된 Geoword 집합과 이를 통해 수집한 웹 정보들로부터 얻은 Non-Geoword 집합, URL 주소 집합들을 서로 분류하고, 이들 사이의 상관관계를 파악하고 순서화 하여 지역성과 상호 연관성을 고려한 지역정보 검색이 가능하도록 하였다. 또한 본 시스템은 Java의 CardLayout 형태의 UI를 응용하고 터치스크린과 스타일러스 펜의 사용을 적극 활용하여 모바일기기의 제한적인 사용자 환경을 극복하였다. 무엇보다 본 시스템은 관광객 등의 모바일 사용자가 자신이 원하는 지역의 지역정보를 검색함에 있어서 기존의 웹 검색 시스템으로는 여러 번의 연관된 정보 검색 절차를 거쳐야 하던 지역정보 검색을 간편하고 효율적으로 할 수 있도록 하였다는 점에 있어 특별한 가치를 부여할 수 있다.

향후 연구로는 특정 값을 지표로 사용하여 수집한 웹 기반의 지역정보를 클러스터링한 후 구조적이고 계층적인 검색이 가능하도록 현 시스템을 발전시키는 것과 지리정보 검색에 랜드마크(Landmark) 요소를 도입하여 보다 사용자 친화적인 검색 시스템을 연구하고 개발하는 것이 있다.

참 고 문 헌

[1] R. Lee, H. Shiina, H. Takakura, Y.J. Kwon, Y. Kambayashi, "Optimization of Geographic Area to a Web Page for Two-Dimensional Range Query Processing", *Web Information Systems Engineering Workshops, 2003. Proceedings. Fourth International Conference on*, pp. 1-9, 13 Dec. 2003

[2] J.H. Jang, R. Lee, Y. Kambayashi, Y.J. Kwon, "Implementation of "Kyonggi21Search" combining GIS with The Web: Optimization of Index Association", *Proc. of The 30th KISS Fall Conference, VOL 30, NUM 2*, pp. 79-81, Seoul, Korea, 24 and 25 Oct. 2003

[3] J.H. Jang, R. Lee, Y. Kambayashi, Y.J. Kwon, "Implementation of "Kyonggi21Search" combining GIS with The Web: index abstraction, index association, user interface", *2003 Symposium on Regional IT Industry Growth*, pp. 19-23, Daejun, Korea, 27 June 2003.

[4] J. Yates and X. Zhou, "Searching the Web using a Map", *Proceedings of the 1st International Conference on Web Information systems Engineering (WISE)*, pp. 222-229, June 2000, Hong Kong.

[5] Java CardLayout, http://java.sun.com/developer/technicalArticles/GUI/