地理學論叢 第31號(1998. 2), 71-90.

GIS를 이용한 질병발생 자료의 지도화와 웹을 통한 공유 방안에 관한 연구

정 재 곤*

The Mapping of Disease Occurrence Data Using GIS and the Sharing of Information through the Web

Jae-Gon Jung

요약 : 본 연구는 보건지리학에서 원자료로 흔히 사용되는 질병발생 자료를 GIS를 이용하여 활용하는 방안을 마련하는 데 그 목적이 있다. 구체적으로 GIS를 이용하여 질병발생 자료를 지도화하고 웹을 통해 다양한 사용자가 효과적으로 공유할 수 있는 방안을 연구하는 것을 목적으로 한다. 질병발생 자료는 위치와 시간을 참조하므로 GIS를 이용하여 분석하고 가시화하는 것은 공간성을 표현할 수 있는 다양한 가능성을 내포하고 있다. 또한 웹 상에서 대화식의 질병발생 지도를 제공하는 경우 사용자가 웹을 통해 질병발생 정보를 좀더 효율적으로 공유할 수 있을 것이다. 본 연구에서는 단순 통계 수치에 그치던 질병발생 자료의 분석 결과물을 GIS를 이용하여 지도화하였으며, 웹 상에서 정적인 질병발생 지도를 제공하는 방식을 발전시켜 대화식으로 제공할 수 있도록 하였다.

주요어: GIS, 보건지리학, 공유, 웹서비스, 질병발생정보시스템, 지도화시스템

Abstract : This study presents the method of GIS application for disease occurrence data which are frequently used as raw data in medical geography. Specifically, it attempts to investigate and show the method of mapping disease occurrence data using GIS and of sharing the information through the web. To analyse and visualize disease occurrence data using GIS has potential benefit since they reference time and location. Furthermore, to provide an interactive map through the web can also realize effective sharing because many web users can communicate with the map. The newly designed system can generate interactive maps from what might be just texts or static maps with past systems. With this system, web users can visually recognize the pattern of disease occurrence.

Key Words: GIS, Medical geography, Sharing, Web service, Disease occurrence information system, Mapping system

Ⅰ. 서론

1. 연구배경

보건지리학(Medical Geography)[1)]에서 주로 사용되는 질병발생 자료는 대부분 위치와 시간을 참조함으로써 시간성과 공간성을 포함한다. 따라서, 질병발생 자료를 기술하고 분석하기 위한 요소로 사람, 시간 및 장소가 도입될 수 있다.

그러나, 지금까지 보건·의료 분야에서는 위치(location) 정보를 고려한 연구가 크게 부족하였다. 따라서, 질병 자료의 효율적인 관리와 정보의 생성에 있어서 GIS를 이용하는 것은 장소의 도입은 물론 공간성을 표현하는 데 있어 다양한 가능성을

* 서울대학교 대학원 석사과정 졸업

제시한다. 특히, GIS를 이용한 지도화는 지도가 전통적으로 중요한 의사소통 도구[2]로 사용되어 왔다는 점에서 보건·의료 분야에서도 크게 활용될 수 있을 것으로 보인다.

질병발생 정보는 체계적인 감시 전산망의 구축을 통해 자료의 수집, 분석을 거쳐 대중 공유를 위해 효율적으로 배분될 수 있어야 한다. 이를 위해 웹(World Wide Web)이 사용될 수 있다. 그러나, 현재 질병발생 정보의 공유를 위해 웹으로 제공되는 것은 간단한 통계적 분석의 결과물, 또는 정적인 지도 이미지[3]나 슬라이드 정도에 그치고 있다. 따라서, 대화식 지도(Interactive map)의 형태로 질병발생 지도를 제공하는 기능이 요구되고 있다. 특히, 클라이언트 쪽에서 동적 디스플레이(Dynamic display) 기능을 구현하는 것은 다양한 질병발생정보 사용자가 웹 상에서 지도와 대화할 수 있는 유용한 수단을 제공한다.

웹을 이용한 질병발생 지도의 제공은 이미 미국, 영국, 프랑스와 같은 선진국에서 효율적으로 이루어지고 있다.[4] 그러나, 국내에서는 현재 질병발생 감시 전산망이 구축 과정에 있으며, 따라서 웹을 이용한 공유도 이루어지지 못하고 있는 실정이다.

2. 연구목적

본 연구는 보건지리학에서 원자료로 자주 사용되는 질병발생 자료를 GIS를 이용하여 활용하는 방안을 마련하는 데 그 목적이 있다.

구체적으로 (1) GIS를 이용하여 질병발생 자료를 지도화하고, (2) 웹을 통해 다양한 사용자가 효과적으로 공유할 수 있는 방안을 연구하는 것을 목적으로 한다.

3. 연구방법

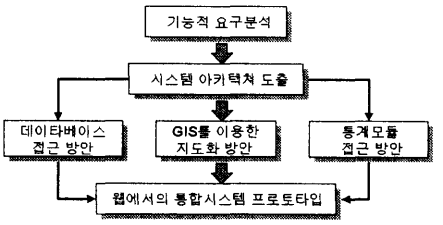
본 연구는 크게 두 부분으로 나눌 수 있다.

1) 원자료로부터 질병발생 정보를 만들어내기 위한 시스템 아키텍쳐를 도출한다. 이를 위해, 분산된 질병발생 데이터베이스, 통계분석도구(Statistical package), 그리고 GIS기능을 가진 모듈 또는 관련 응용프로그램에 접근하기 위한 메카니즘을 정의한다.
2) 질병발생 자료를 지도화하고 웹 브라우저로 제공할 수 있는 웹 인터페이스를 설계 및 구현한다. 지도화를 위해 자료의 속성에 맞는 심볼을 선택하여 구현하고, 사용자가 대화식으로 처리할 수 있는 기능을 추가한다.

이를 통해, GIS를 이용하여 질병발생 자료의 분석 및 공유 기능을 담당하는 통합 시스템 프로토타입(Prototype)을 웹 상에서 구현한다. 이를 위해, 부가적으로 웹-DBMS 인터페이스와 웹-통계모듈 인터페이스를 설계 및 구현한다.

4. 연구 흐름도

연구 흐름도는 [그림 1-1]과 같다.
먼저 기능적 요구분석에 기반하여 시스템 아키텍쳐(System architecture)를 도출하며, 이를 바탕으로 GIS를 이용한 지도화 방안을 마련한다. 부가적으로 데이터베이스와 통계모듈로의 접근 방안이 마련됨으로써 통합 시스템의 프로토타입을 구현한다.



〈그림 1〉 연구 흐름도

5. 자료와 사용 시스템

시스템의 구현을 위해 사용된 주요 자료는 서울

을 기본 연구지역(region)으로 하고 있다. 이는 서울지역에서 수집되는 질병발생건수는 공간 단위별로 다른지역에 비해 많으므로 프로토타입의 구현이 효과적이라고 판단되었기 때문이다.

이용된 자료는 상병(傷病) 자료, 대기오염 자료 및 인구 자료이며, 오라클(ORACLE) 데이터베이스에 저장되었다. 상병자료는 비감염성 질환에 해당하는 질병[5]들로, 1992-1995년 사이의 의료보험 자료를 사용하였다. 비감염성 질환의 환경적 요인 중의 하나로 중요하게 인식되는 대기오염 자료는 1993-1995년 사이의 자료를 이용하였다. 인구 자료는 센서스 자료에 기반한 연앙 추계 인구를 사용하였다.

통계적 분석을 위해서는 S-minus, 서버쪽의 기능을 구현하는 GIS 시스템으로는 GRASS를 이용하였다.

Ⅱ. 이론적 고찰과 감시전산망 구축현황

1. 질병발생 자료와 GIS

1) 보건지리학의 개념 및 연구 동향

보건지리학은 인문지리학의 한 분야로 인간의 건강과 보건이 갖는 지리적 측면을 연구한다.

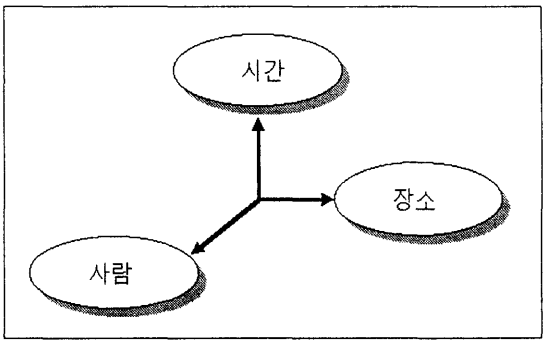
역사적으로 사람들은 어떤 질병이 특정 장소에서만 발생하고 다른 곳에서는 발생하지 않는다는 사실을 경험하고 있었다. 즉, 건강에 영향을 미치는 결정요인으로서 장소가 가지는 중요성을 일찍부터 인식하고 있었다.[6]

지리적 영역에서 질병 확산을 지도화하기 시작한 것은 1854년 스노우(John Snow)에 의해서이다. 그는 런던의 소호(Soho) 지역에서 발생한 콜레라에 대한 연구에서 콜레라 발생을 지도화하였다. 그는 지도화를 통해 질병발생의 원인을 설명하고, 질병 확산을 막기 위해 우물을 폐쇄하는 근거를 마련할 수 있음을 보여주었다.[7] 19세기에 지리학이 학문으로 성숙하면서 지리학자들이 보건과 질병이 갖는 지리적 측면에 관심을 보이기 시작하였으며, 이후 활발한 연구를 통해 보건지리학이 지리학의 한 분야로 정착되기 시작했다.

최근 연구에서는 질병분포를 가시화하기 위해 자료를 지도화하는 경향이 두드러진다. 특히, 정적인 이미지(Static image)들을 슬라이드의 유형으로 보여줌으로써 질병확산을 인식할 수 있도록 도와주는 방식을 자주 사용한다.[8]

2) 보건 지리학에서 GIS의 필요성과 역할

인간의 건강은 사회문화적, 물리적 환경을 포함한 환경적요인에 의해 크게 영향을 받으므로 항상 환경적, 공간적 차원을 내재하고 있다. 따라서, 질병발생 자료를 기술하고 분석하기 위한 요소로 사람, 시간 및 장소가 도입될 수 있다(Raymond et al, 1995 : pp. 28-29).[9] 이 중에서 특히 장소는 질병발생의 주요 요인 중 하나이다.



자료: Raymond et al, 1995, p. 29.

〈그림 2-3〉 질병발생을 특징짓는 개념적 표현

그러나, 지금까지 보건·의료 분야에서의 질병발생 연구는 장소의 관점에서 바라보는 경우가 극히 드물었다. 이러한 한계는 질병발생 자료의 분석 및 가시화에 있어서 GIS를 활용하는 것이 공간성을 포함하는 영역으로 기존의 연구를 확장시킬 수 있다는 가능성을 보여준다.

1990년대 들어 GIS는 값싸고 강력한 지리적 분석 도구가 되었다. 비용의 감소와 더불어 GIS가 가지는

공간분석과 가시화의 장점은 이를 이용한 보건 지리학적 연구[10]가 폭발적으로 증가하도록 만들었다.

전통적으로 GIS 기술은 그 적용 범위를 다양한 분야로 확대시켜 왔으며, 통합적(integrative)이면서도 다학문적(interdisciplinary)인 해결방안을 제시하려 하였다(Michell-Viret, 1994).

그러나, 농업, 자연자원, 도시, 지역계획, 관광 분야에서 이미 일찍부터 GIS를 이용해 온 것과 달리 보건지리학 및 타 보건학 관련 분야는 최근에서야 GIS적 분석을 도입하고 있는 상황이다.[11] 따라서, GIS의 응용 분야를 보건학과 보건 지리학으로 확장하는 것은 다학문적 연구와 보건 시스템의 개발에 힘입어 가장 새롭고 흥미로운 분야가 되고 있다.

GIS는 질병 위험요소의 분포와 상호작용, 질병발생과 사망의 공간적 패턴, 그리고 보건 자원의 분배에 관한 좋은 정보를 제공할 수 있다. 따라서, 이를 통해 보건 연구, 보건 교육, 계획, 모니터링과 감시, 보건 계획의 평가 등에 유용하게 사용될 수 있다.

3) GIS를 이용한 질병발생 자료 처리

보건지리학에서 다루는 많은 자료 중에서 질병발생 자료는 환경오염이나 유전적 요소 등의 질병발생 요인들(factors)이 만들어내는 결과물이라는 점에서 매우 중요한 의미를 가진다. 즉, 질병발생 자료를 분석함으로써 질병발생 요인들의 영향 정도를 평가하고, 시공간적 변화 양상을 바탕으로 앞으로의 질병발생을 예측할 수 있다는 점이 중요하게 인식된다.

이러한 질병발생 자료의 중요성은 공간적 분포 양상 및 시간적 변화 양상을 파악하는 것이 질병의 예측과 예방에 큰 역할을 할 수 있음을 암시한다. 따라서, 공간성과 시간성을 효과적으로 분석하고 가시화할 수 있는 GIS 시스템의 사용이 요구되고 있는 실정이다.

2. 질병발생 정보의 대중 공유

1) 공용 정보로서의 질병발생 정보 공유

지리정보의 공유에 관한 연구는 비용절감에 따른 이익, 기관간의 협력, 공유자료 수집의 메카니즘, 자동화, 저장, 그리고 효율적 사용 등 실용적인 문제에 초점을 맞추어왔다. 그러나, 최근들어 공간정보에 기반한 GIS 데이타를 일반 국민 및 지리적으로 분산된 일반 분석가들과 공유하는 문제가 좀 더 중요하게 부각되고 있다(Arne, et al., 1997). 질병발생 정보는 국민의 건강과 관련된 공용정보(public information)로서 일반 대중에게 다양한 유형으로 제공될 필요가 있다. 이를 통해 일반 대중이 질병발생 정보에 직접 접근할 수 있는 기회가 많아질수록 공중 보건의 주요 요인이 되는 환경에 대한 보호와 보건 서비스의 확충이 더 잘 이루어질 수 있을 것으로 기대된다(Qi, et al., 1997). 더욱이 공간성을 기반으로 표현된 질병발생정보는 일반인이 공간에 대한 인식을 확장시키는 데 도움을 줄 수 있다.

현재 가장 일반적인 공간 정보 공유 도구는 지도이며, 이를 기초로 일반사용자는 다른 유형의 데이타를 추출할 수 있다. 그러나, 우리나라의 경우, 공간정보를 담고 있는 질병발생 정보에 대한 접근 도구로서 질병분포를 보여주는 종이 지도조차도 폭넓게 이용되지 못하고 있으며, 일반적으로는 단순한 통계 수치를 제공하는 데 그치고 있다.[12]

2) 질병발생 정보 공유의 기술적 제한점

질병발생의 공간적 요소에 대한 공용접근은 앞에서 언급했듯이 일반 국민이 질병발생 분포 및 변화 양상에 관한 공간적 인식을 확장시키는 데 도움을 줄 수 있다는 장점이 있으나, 기술적 문제13)로 인하여 활발하게 이루어지지 못하고 있다. 따라서, 공용접근의 기술 개발을 통해 공유 능력을 향상시키는 것이 중요한 요소가 된다. 인터넷과 웹 수요의 폭발적인 증가는 이를 이용한 질병발생 정보의 제공을 더욱 현실화시키고 있다. 따라서, 앞으로 이를 포함한 타분야의 정보기술 개발요소들이 현재의

GIS 기술에 빠른 속도로 통합할 수 있어야 질병발생 정보의 효과적인 공유가 쉽게 구현가능할 것으로 보인다.(Qi et al, 1997 : p556).

3. 질병발생 감시 전산망 구축

1) 질병발생에 대한 시스템 접근 방식

연구자들이 어떤 문제를 해결하기 위해 적합한 모델링 도구를 사용할 수 있도록 도와주거나 일반 사용자나 정책 결정자들에게 정보를 제공하기 위해서는 효율적인 질병발생 감시 전산망을 구축하는 것이 바람직하다. 이를 위해서는 질병발생의 성격에 따른 공간의학적 문제들(Geomedical problems)의 규명이 먼저 이루어져야 한다. 왜냐하면, 질병발생의 효과적인 관리 및 예방은 각각의 질병 유형에 맞는 적절한 감시 시스템의 구현을 통해 가능하기 때문이다. 시스템 접근방식을 이용하여 공간의학적인 문제들을 세 가지로 구분하면 다음과 같다(Richards, 1992).

첫째, 감염성 질병을 위한 전염병 예측 시스템.
둘째, 비감염성 질병을 평가하기 위한 통계 시스템.
셋째, 보건 서비스 시스템을 위한 사회-경제학적 모델.

이러한 접근방식을 이용하는 것은 구축될 감시시스템의 유형 및 기능을 파악하는 데 유용하다.

2) GIS와 분산 자원의 관리 방안

정보사회는 데이타를 기초로 하며, 이는 가공되지 않은 자원이라 할 수 있다. 자동화와 정보사회의 본질은 데이타의 생성, 처리, 조작을 통해 정보를 제공하는 것으로 이러한 정보 제공의 기능은 분산 시스템에 의해 보강될 수 있다(채기준 외, 1995). 자원을 분산시켜 관리하는 것은 질병발생정보가 가지는 공간적 요소의 경우에도 예외는 아니다. 이 경우에 분산 시스템의 제어나 공용 서비스 등과 같은 여러 가지 문제를 해결하기 위해 분산 기술을 사용하여 구축하는 경향을 보이고 있다. 분산 GIS 기술을 사용하는 경우 미들웨어 개발을 통해 복잡한 구조를 단순화시키고, 앞으로의 분산 기술에 대한 호환 및 확장성을 보장할 수 있다.

3) 웹을 이용한 질병발생 정보 제공

폭발적으로 증가하고 있는 웹 수요와 기술적 진보로 인해 거의 모든 관련 기술이 영향을 받았다. 공간정보기술도 예외는 아니어서 이미 다양한 유형의 공간정보가 웹에서 사용가능한 수준까지 이르렀으며, 기존 상용 GIS 프로그램의 기능을 웹에서 사용할 수 있도록 하는 노력이 계속되고 있다.

인터넷과 웹 모두 공간 정보를 공유하기 위한 기반시설을 제공한다. 인터넷은 전세계에 분포하고 있는 컴퓨터와 통신망을 하나로 연결시킨 것으로 유용한 응용프로그램을 사용할 수 있도록 하며 웹(World Wide Web) 은 하이퍼미디어를 제공할 수 있는 또 다른 유용한 응용프로그램으로 인터넷에 연결된 컴퓨터의 자원을 쉽게 접근해서 쓸 수 있도록 한다. 현재 웹은 풍부한 텍스트와 이미지, 소리, 동영상과 같은 하이퍼미디어(Hypermedia)를 인터넷에서 사용가능하게 만드는 응용프로그램으로 인식되고 있다(Blaine, 1997 : p. 547).

웹 기술을 이용하면 단순한 데이터 전달 이외의 기능을 보완할 수 있으며, 질병발생 정보의 공유를 위해 웹 요소들의 통합 방안이 고려될 수 있다.

4. 우리나라의 질병발생 감시 전산망 구축 현황

현재 우리나라는 질병발생 감시 전산망의 구축 과정에 있다. 최근들어 활발하게 논의되고 있는 질병발생 감시체계(Surveillance system)[14]의 구축은 크게 감염성 질환정보의 수집 및 배분 분야[15]와 비감염성 질환의 발생 원인으로서 환경적 요인의 영향을 분석하는 분야[16]로 나누어볼 수 있다.

특성별로 구축이 가능한 각 시스템들은 공통된 기능을 요구하는 경우가 많으므로 공통의 시스템

구성요소들을 통합하는 방식이 고려될 수 있다.

이러한 측면들을 고려할 때, 현재 효과적인 분석과 공유를 위한 감시 전산망의 구축이 시급한 실정이다.

Ⅲ. 지도화 시스템 설계

1. 시스템의 기능적 요구분석

1) 분산 처리

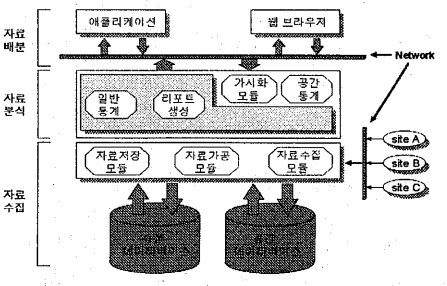
질병발생과 관련된 자료는 다양한 유형의 시스템에 분할되어 관리되는 경우가 많다. 따라서, 질병발생 관련 자료들이 저장될 수 있는 이질적 시스템(heterogeneous system)에 통합적으로 접근하기 위해서는 클라이언트/서버 방식이 기본적으로 고려되어야 한다.

질병발생정보는 다양한 데이터베이스에 저장된 정보를 추출하여 GIS 시스템에서 다양한 공간분석 및 공간 가시화 도구를 통해 만들어질 수 있다. 따라서, 데이터 통합을 통해 다양한 유형의 데이터베이스에 자유롭게 접근할 수 있으며, 응용프로그램의 통합을 통하여 질병 발생 정보의 생성에 필요한 여러 가지 응용프로그램들을 효율적으로 이용할 수 있다.

2) 질병발생 감시 전산망에서의 배분 기능

질병발생 감시 전산망은 질병발생 자료의 수집, 분석, 배분과 관련된 다양한 기능을 하나로 통합하는 것이어야 한다. 따라서, 각각의 기능을 갖춘 시스템은 하나의 하부시스템(subsystem)으로서 전체 질병발생 감시 시스템에 통합될 수 있어야 한다.

〈그림 3-1〉은 질병발생 감시시스템을 자료 수집, 분석 및 배분의 관점에서 구성한 것이다. 여러 지역(site)에서 수집된 질병발생 자료는 필요한 가공을 거쳐 공간 및 비공간 데이터베이스에 저장되고 (① 자료 수집 단계), 일반통계 및 공간통계와 같은 분석 과정을 거치게 되며 (② 자료 분석 단계), 지도화와 같은 가시화 과정을 거쳐 사용자에게 제공된다. (③ 자료 배분 단계).

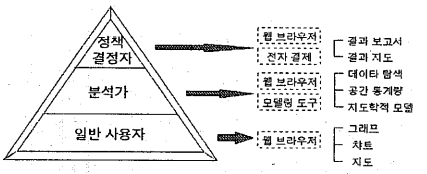


〈그림 3-1〉 감시 시스템 기능의 개념적 구성

질병발생 자료의 지도화를 통한 웹에서의 제공은 질병발생 감시 전산망의 기능 중에서 자료의 분석 및 결과물의 배분 기능과 밀접하게 관련되어 있다고 볼 수 있다. 따라서, 시스템의 기능적 요소들은 데이터의 공간통계적 분석과 가시화를 통한 웹 서비스에 중점을 두어야 한다.

3) 사용자 중심의 대화식 지도생성

질병발생 정보의 사용자는 각각의 수준[17]에 따라 그 수가 다르며, 사용자 수에 의해 [그림 3-2]와 같이 피라미드 형태로 개념화될 수 있다.[18] 사용자들이 원하는 정보의 형태 및 정보 획득의 방식은 차이를 보일 수 있다. 그러나, 웹이 가지는 범용성으로 인해 일반적으로 웹 브라우저를 통한 접근 방식은 세 종류의 수요 집단에 공통적으로 적용될 수 있다.



〈그림 3-2〉 질병발생 정보 시스템 사용자의 개념적 분류

사용자의 지식이나 정보의 이용 목적 등이 다르다는 사실은 웹 인터페이스의 구현에 있어서 다양한 사용자의 요구를 충족시켜야 함을 의미한다. 이러한 사용자의 요구는 대화식 기능을 추가함으로

써 해결할 수 있다. 그러나, 아직 웹 상에서 제공되고 있는 공간정보를 위해 대화식 기능을 지원하고 있는 경우는 거의 없다.

사용자가 가시화 시스템을 얼마나 잘 제어할 수 있는가를 결정하는 사용자 제어(user control) 특성들로는 분산, 대화식 제어, 후단처리, 모듈화, 통합 등의 요소가 있다(Hilary et al, 1994). 이 요소들은 사용자와의 친숙도를 결정하므로 시스템의 설계 및 구현에 고려해야 한다.

웹 브라우저를 이용한 질병발생 자료 및 정보의 공유는 폐쇄된 자료 집계 및 분석 과정을 개방된 구조로 변화시킬 수 있다는 또다른 장점을 지니고 있다.

즉, 지금까지의 질병발생 정보의 공유는 일반 사용자의 입장에서 고려되었지만, 분석가의 입장에서 질병발생 자료를 공유하는 것은 더많은 가능성을 내포한다.

2. 웹을 통한 공유 시스템 구조 설계시 기술적 고려사항

1) 웹 인터페이스의 유형

대화식 지도화(Interactive Mapping)에 있어서 가장 큰 논란이 이는 주제 중의 하나가 GIS-웹 인터페이스의 유형이다.(Bill, 1996) 특히, 웹에서 사용되는 그래픽 사용자 인터페이스(Graphic User Interface)는 서버쪽에서 제공하는 정보의 종류나 유형에 의존하므로 어떤 유형의 정보를 서버쪽에서 제공하는가가 문제가 된다. 이를 해결하는 방식으로 GIS 분야의 정보 유형에 맞춘 일반적인 인터페이스를 제공하는 것이 고려될 수 있다.

2) 웹 이용(Web-enabling) 기술의 비교

① CGI 방식과 브라우저 포함 응용프로그램 방식

구현 가능한 유형 중에서 가장 많이 사용되는 방식으로는 폼에 기반한 CGI(form-based page using Common Gateway Interface) 방식과 자바(Java)나 플러그-인(Plug-in) 프로그램처럼 브라우저에 포함시킬 수 있는 응용 프로그램에 의한 방식이 있다. 브라우저 포함 응용프로그램을 이용하여 GIS 시스템과 연결시키는 경우에는 진정한 대화식(Interactive) 인터페이스를 구현할 수 있다는 장점이 있으나, 반면에 전문적인 구현을 위해서는 숙련된 프로그래밍 기술이 필요하며, 실행속도가 느리다는 단점이 있다. CGI를 이용한 방식은 좀더 빠르고 믿을만 하지만 완전한 대화식 환경으로의 확장에 문제가 있다. 따라서, 정보의 유형에 따라 이 둘을 혼용하는 방식(Hybrid Approach) 이 사용될 수 있다.

② 자바 애플릿과 티클 플러그인

대화식 지도를 웹 상에서 제공하기 위해 사용되는 방식 중에 가장 많이 사용되는 방식 중의 하나가 CGI와 자바를 이용하는 것이다. 자바는 단순하고, 객체지향적이며, 분산 환경에 적합하고, 인터프리터 방식이며, 견고하고, 안전하며, 아키텍쳐 중립적이고, 이식성이 좋으며, 높은 성능을 가진 동적인 언어이다.[19]

현재, GIS에서 사용하는 공간정보를 웹에서 대화식으로 제공하는 웹 페이지는 거의가 자바를 구현 언어로 채택하고 있으며, 웹 언어로서 자바가 가지는 가능성은 앞으로의 확장에 있어서도 큰 장점을 가질 것으로 예상된다. 티클(Tcl)[20]은 유닉스 기반의 스크립트 언어로 윈도우 환경에서의 프로그래밍 환경을 제공한다(Jason, 1996 : pp. 178-179). 웹 브라우저 상에서 티클 프로그램을 사용하기 위해 플러그인(plug-in)이 개발되었으며, 이를 이용하여 기존의 프로그램을 인터넷 상의 사용자들에게 제공할 수 있다.

본 연구에서는 티클이 스크립트 언어로서 가지는 장점, 즉 프로토타입의 개발이 용이하다는 점을 크게 고려하였으며, 따라서 CGI 와 티클 플러그인을 이용하여 대화식 지도를 제공하였고, 필요한 경우에는 CGI와 자바를 이용하는 방식을 추가적으로 취할 수 있다.

3) 접근 구현의 방식

기존의 응용프로그램을 웹에서 이용(Web-enabling)할 수 있도록 만드는 방식으로는 크게 변환(translation)[21], 공간 엔진의 재설계, 웹 언어에 공간 연산자를 추가하는 방식의 세 가지를 들 수 있다. 현재 변환 방식이 가장 일반적으로 쓰이고 있으며, 이는 웹에서 사용하는 HTML(Hypertext Markup Language : Lemay, 1995)과 CGI(Common Gateway Interface), 자바 애플릿(Java Applet), 플러그인(Plug-in) 기술 등 진보된 웹 기술에 의해 가능하다.

본 연구에서는 변환 방식을 이용하였다. 이는 성능보다는 프로토타입의 개발에 중점을 두었기 때문이며, 대화식 환경은 서버와의 유기적 연결보다는 클라이언트 쪽(client-side)에서 지원하는 방식을 택하였다.

4) 시스템 구조 설계 방식

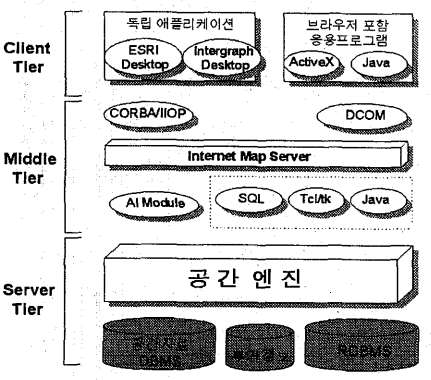
시스템 구조의 설계에 있어 고려해야 할 기술로는 클라이언트, 미들웨어, 그리고 서버 기술을 들 수 있다.

클라이언트로는 독립적 애플리케이션(standalone application) 이나 브라우저 기반의 애플리케이션(browser based application) 이 쓰일 수 있으며, 각각의 경우에 적합한 서버 연결 모듈이 개발되어야 한다. 클라이언트 기술은 사용자의 요구를 받기 위한 GUI(Graphic User Interface) 설계에 치중되어 있다.

미들웨어는 클라이언트와 서버간의 연결을 유동적(flexible)으로 만들기 위해 사용될 수 있다. 클라이언트의 요구를 다양한 방식으로 자동화하기 위해 인공지능(Artificial Intelligence) 기술이 이용되거나 분산기술로 CORBA/IIOP, DCOM 등이 사용될 수 있다.

서버 기술은 변환 방식을 응용하는 경우 클라이언트와 서버쪽의 프로그램을 연결할 수 있는 모듈 개발이 주를 이루며, 공간 엔진을 재설계하는 경우에는 공간 엔진 자체가 클라이언트의 요구를 처리할 수 있다.

분산 환경에 적절하게 대응하기 위해서는 3-티어(tier) 아키텍처를 사용함으로써 시스템의 확장성을 높일 수 있다. 이는 시스템의 기능을 3개의 티어로 나누고 그 각각에 맞는 구성 요소들이 역할을 분담하는 구조라고 할 수 있다. 〈그림 3-4〉는 3-티어 아키텍처를 GIS 시스템 설계에 적용할 때 제시될 수 있다.



자료: 97' GIS Workshop proceedings, cadland, 참조.

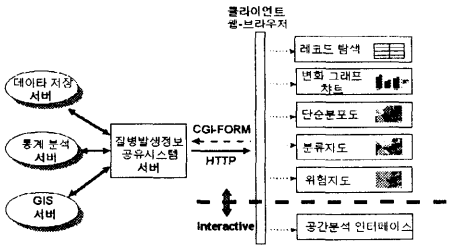
〈그림 3-4〉 3 티어 아키텍처

3. 지도화 시스템 아키텍처

1) 기능과 구성요소

지도화 시스템에서는 자료의 효율적 가시화 및 배분이 분석적 작업과 연계되어야 한다. 따라서, 사용자의 요구는 CGI를 이용하여 서버에 전달되고, 서버에서의 처리를 통한 결과물은 각각의 유형을 가장 잘 전달할 수 있는 기능을 포함한 형태로 제공되는 것이 좋다.

질병발생 정보 공유시스템 서버는 가장 핵심이 되는 서버의 기능을 담당하며, 사용자 요구를 받아서 어떤 시스템에서 처리할 것인지를 결정한다. 사용자가 선택한 처리 과정을 거치면 그 결과물은 기초적 데이터 탐색을 위한 레코드 검색 및 그래프 검색, 그리고 지도화를 거쳐 웹 브라우저 상에 보여진다.



〈그림 3-5〉 질병발생정보 시스템의 전체 개념도

서버와 클라이언트와의 관계는 대화식인지 아니면 비대화식인지에 따라 크게 두 가지 유형으로 나누어질 수 있다.

대화식인 경우에는 클라이언트의 요구가 실시간으로 서버에 전달되고 그 결과물을 받아온 후에도 필요한 경우 연결을 지속한다. 반면에 비대화식인 경우에는 한 번 클라이언트에 전달된 프로그램은 더 이상 서버와 연결하지 않는다.

본 연구에서 설계한 지도화 시스템은 비대화식이며, 보완적으로 추가된 그래프와 웹-통계 모듈 인터페이스 중에서 웹-통계 모듈 인터페이스에서의 클라이언트만이 대화식으로 서버에 접근할 수 있도록 하였다.

즉, 웹-통계 모듈 인터페이스는 완전한 대화식 접근을 지원하며, 따라서 웹 브라우저를 클라이언트로 한 서버와의 대화가 가능하다.

2) 서버 아키텍쳐 도출

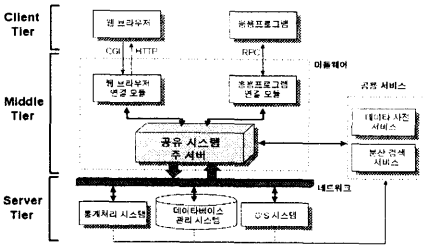
① 미들웨어로서의 서버

미들웨어는 개념적으로 여러 개의 클라이언트와 여러 개의 서버를 중간에서 연결하는 기능을 담당하는 하나의 또 다른 서버를 이용하여 구현될 수 있다. 이것은 응용프로그램이나 웹 브라우저가 접근을 시도할 경우 전달하는 정보를 받아서 각 기능별 서버에 다시 전달하고 결과를 받아 넘겨주는 역할을 한다.

〈그림 3-6〉을 보면, 시스템 구조가 3-티어로 되어 있음을 알 수 있다. 미들웨어에 포함되는 주 서버는 네트워크 상의 기능별 서버들에 접속한 후 원하는 결과물을 클라이언트에 건네준다. 각 기능별 서버들이 서비스를 시작하면 그 정보가 공용 서비스를 담당하는 서버에 등록될 수 있다.

② 클라이언트로부터의 접근

클라이언트는 크게 웹 브라우저를 사용하는 경우와 독립적인 애플리케이션을 사용하는 경우로 나누어볼 수 있다. 현재의 웹 기술을 살펴보면, 애플리케이션을 제공하는 경우에는 사용가능한 자원(Software Resource)이 많다는 장점이 있는 반면에 웹 브라우저 상의 애플릿을 사용하는 경우에는 일반인의 접근이 쉽다는 장점이 있다. 두 방식이 가지는 장단점으로 인하여 두 가지 방식을 유연하게 통합할 수 있는 구조를 제시할 수 있다.



〈그림 3-6〉 미들웨어로서의 기능을 하는 서버 개념도

Ⅳ. 질병발생자료의 지도화와 웹에서의 공유 시스템 구현

1. 면 심볼을 이용한 지도화와 웹에서의 공유

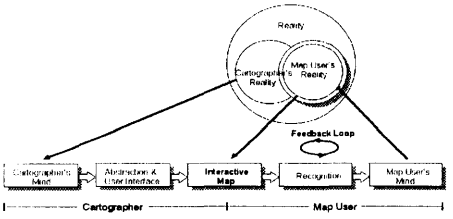
1) 질병발생 자료의 지도화

GIS에서는 조건분석(What if? analysis)의 결과를 전달해주기 위한 수단으로 지도를 주로 사용한다(Hilary et al, 1994). 지도학적 가시화(cartographic visualization)는 때로 지리적 가시화(geographic visualization)로 불리며, 지도 디스플레이와 유사한 기

술의 사용으로 해석되는데, 이에 관한 연구는 일반적인 모델의 도출이 가능하게 만들었다(Michael, 1995).

이 관점에서 보면 지도학적 의사소통은 정보가 선택되고 지도위에 심볼로 표현되며, 지도 이용자에게 인식되고 해석되는 과정으로 정의된다. 이 과정은 크게 두 개의 분리된 영역, 즉 지도 제작자의 영역(cartographer's domain)과 지도 사용자의 영역(map user's domain)으로 나누어진다. 지도 제작자(cartographer)의 역할은 지도 이용자가 인식하는 현실세계와 일치하는 부분들을 설명하기 위해 지도를 만들어내면서 의사소통 과정을 제어하는 데 있다. 이 과정에서 현실세계는 지도제작자에 의해 추상화(abstraction)된다.

전통적으로 지도학적 의사소통의 개념은 평균적 지도 사용자(average map user)의 개념에 기반하고 있었으며, 따라서 그러한 사용자를 대상으로 만들어지는 정적이고 표준적인 지도가 대부분이었다.



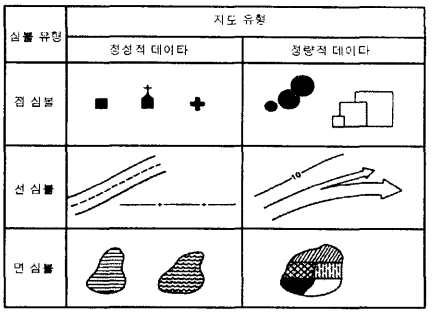
자료: Michael, 1995 : P. 6

〈그림 4-1〉 지도학적 의사소통 모델

그러나, 지도 사용자가 디스플레이된 지도와 대화할 수 있는 가능성을 더 많이 가지는 경우 정적인 지도에서 가졌던 한계를 극복할 수 있다. 대화식 지도(interactive map)는 피드백(feedback loop)의 결과물이며, 피드백은 사용자 인터페이스 안에서 구현되는 반응 메카니즘으로 이루어진다.

2) 다양한 심볼의 사용

일반적인 참조(reference)와 주제도에 공통으로 사용될 수 있는 세 가지 기본 유형의 심볼로 점(point), 선(line), 면(area)을 들 수 있다. 데이터의 특성에 따라 서로다른 심볼을 사용하여 표현할 수 있으며, 같은 유형의 공간 참조 데이터라고 할지라도 다른 심볼을 사용하여 다양한 유형으로 표현할 수 있다. 행정구역과 같이 기본 공간객체를 면으로 하여 집계(aggregation)된 질병발생 자료는 양적(quantitative)인 것이므로 면 심볼 중에서 양적인 자료를 표현할 수 있는 것이어야 한다.



자료원: Borden, 1993, p. 84.

〈그림 4-2〉 다양한 지도 심볼과 지도 유형,

① 점 심볼을 이용한 질병발생 자료의 지도화

점 지도(dot maps)는 점 심볼을 이용하며 점 개개의 질병발생 자료가 가지는 위치를 참조하여 공간적 분포를 보여준다. 질병발생수의 증가는 점의 수 증가와 비례하여 표현되며, 도시 기반 시설이나 보건 서비스 시설의 위치와 같이 여러 가지 다른 벡터 정보를 동시에 보여줌으로써 데이터 위치의 중첩(overlay) 기능을 보여줄 수 있다. 이것은 질병발생 지도화의 가장 단순한 방법으로 질병 발생지역의 표현에 유용하게 사용될 수 있다.[22]

GIS에서 쉽게 구현될 수 있는 것으로 점 지도에 동적 그래픽 기법을 사용하는 방법이 있다(Haslett, 1991). 이 접근 방법을 사용하면 각각의 질병 발생 지점은 히스토그램의 질병 발생건수와 내부적으로 연결되며, 따라서 특정 히스토그램을 선택하면 그것에 해당하는 발생 위치가 지도에 자동으로 표시된다. 이 방식을 이용하여 질병 발생을 분류하고

질병 발생율이 높은 몇몇의 지역을 쉽게 확인할 수 있다. 질병발생 자료는 일반적으로 집계(aggregation)되어 분석되는 특성을 가지고 있는데 이 경우에는 주로 면 심볼을 사용하며, 개개의 질병발생 자료를 직접 사용함으로써 점 심볼을 사용하는 것은 주로 소규모 지역(micro-scale)에서의 깊이있는 연구에 유용하다.

② 선 심볼을 이용한 질병발생 자료의 지도화

벡터나 선은 질병의 확산과 환자-보건 서비스 시설간의 흐름을 분석하는 데 도움을 줄 수 있는 그래픽 자원이다. 가장 단순한 유형으로 선은 두 개의 하부지역 간에 전염의 흐름이 존재하는가를 확인시켜 준다. 화살표와 선의 두께 또는 선에 특정 자료값을 할당함으로써 다양한 흐름을 표현할 수 있다. 선 심볼을 이용한 질병발생 자료의 지도화는 에이즈와 같은 몇몇 전염성 질환의 확산을 전지구적 스케일에서 표현하는 데 자주 이용된다.

③ 면 심볼을 이용한 질병발생 자료의 지도화

기본 공간객체를 면으로 하여 집계된 자료를 지도화한 경우 흔히 코로플레스 맵(choropleth map)[23]으로 불리며, 가장 자주 사용되는 방식 중의 하나이다.[24]

면 심볼을 이용한 질병발생 자료의 지도화에 사용되는 주요 변수로는 발생률(incidence rates), 사망률(mortality rates), 표준화된 사망률(SMR : Standardized Mortality Ratio) 등이 있다. 이 중에서 표준화된 사망률 지도는 질병분포지도(disease distribution map) 중에서 가장 일반적인 유형이다(Walter et al. 1991).

면을 기본 공간객체로 하는 지도화는 점을 기본 공간객체로 하여 수집되는 자료보다 통계적으로 유의한 자료를 사용하며, 때로는 상대적 위험도에 대한 경험적 베이즈 측정치(empirical bayes estimates)를 사용한다(Marshall, 1991). 각 공간 기본단위(spatial unit)가 두 개 또는 그 이상의 변수를 가지고 있을 경우에는 다변량 분석(multivariate analysis)을 통해 정보를 합성할 수 있다. 최근에는 두 변수간의 상관성(coherency)을 이용하여 질병 발생의 공간적 패턴을 설명하기도 하였다(Cliff et al, 1992).

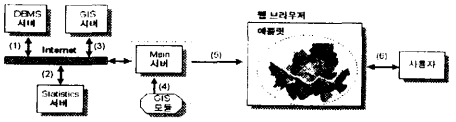
코로플레스 맵의 또다른 문제는 행정 경계를 기본 단위로 사용하는 데서 발생하며,[25] 이를 해결하기 위해 지도변환(map transformation)26) 등의 방법이 이용된다.

3) 지도화 시스템의 연결구조

지도화 시스템이 사용하는 서버는 데이터베이스 서버와 통계분석 서버 그리고 GIS 서버이다. 데이터베이스 서버로부터는 원하는 조건의 질병발생 자료를 검색(query) 및 집계(aggregation) 한 후에 가져오며(〈그림 4-3〉의 1번), 통계분석 서버로부터는 검색된 자료를 원하는 유형으로 통계분석한 결과물을(〈그림 4-3〉의 2번), 그리고 GIS 서버로부터는 질병발생 자료를 구를 기본 공간단위로 하여 다시 집계한 후 보간(Interpolation)하고 등치선 지도를 만들어서 가져온다(〈그림 4-3〉의 3번).

미들웨어의 기능을 하는 주 서버는 이러한 일련의 작업을 총괄하는 역할을 담당한다. 즉, 클라이언트로서의 웹 브라우저는 주 서버에만 접속한 후에 원하는 메시지를 전달하기만 하면 된다. 주 서버에서는 GIS 서버의 기능을 직접 사용하지 않는 경우, 공간좌표를 할당하고(〈그림 4-3〉 의 4번) 면 심볼로 표현한 후에 웹 브라우저로 전송한다(〈그림 4-3〉의 5번). 이 과정에서 질병발생 자료를 통계적으로 표현할 수 있는 다양한 기능들이 주 서버에 의해 포함된다.

웹 브라우저 상에 브라우저 포함 응용프로그램의 형태로 들어가 있는 지도화 시스템은 사용자의 다양한 요구를 대화식으로 처리해 줄 수 있는 기능을 가지고 있다(〈그림 4-3〉의 6번). 즉, 웹 브라우저 상에 포함된 애플릿 유형의 결과물은 가시화 기법들을 이용하였으므로 사용자가 원하는 유형의 결과물을 직접 조작하여 탐색할 수 있다.



〈그림 4-3〉 지도화 시스템 아키텍쳐

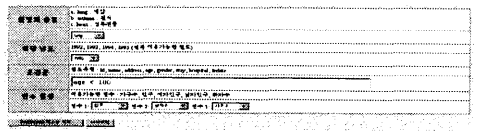
4) 표현방식 및 기능

① 코로플레스 맵의 생성

개개의 질병발생 자료는 개인정보를 포함하고 있다. 따라서, 기밀성(confidentiality)을 유지해야 하므로 공간 기본단위를 점으로 하는 자료에 접근하는 것이 대부분의 경우 불가능하다. 대신에 구/ 시/ 군이나 읍/ 면/ 동 경계와 같은 공간단위별로 집계된 자료를 사용하는 것이 일반적이다. 지도화 시스템은 기본적으로 코로플레스 맵을 지원함으로써 동을 기본 공간단위로 하는 자료의 가시화에 중점을 두었다.[27] 〈그림 4-4〉는 지도화 시스템을 시작하기 위한 CGI 폼으로 질병의 종류와 연도의 선택, 조건문 입력 및 변수 설정이 가능하다. 데이터베이스에 저장된 질병 자료에 따라 CGI 폼에서 지원하는 질병의 종류가 정해지게 되므로 또다른 종류의 질병을 데이터베이스에 저장하는 경우 폼의 갱신만으로 그 자료를 처리할 수 있다. 처리할 수 있는 시간 스케일(time scale)은 1년이며, 데이터베이스에 저장된 자료의 시간필드(time field)가 1992년부터 1995년까지이므로 4가지 선택이 가능하다. 시간 스케일을 한 달 또는 2 내지 3년으로 정할 수 있는 기능은 필요한 경우 새로운 모듈의 추가를 통해 구현할 수 있다. 조건문 필드에는 데이터베이스 처리에 있어 SQL 의 where 문을 입력할 수 있으며, 아무런 제한없이 조건문이 데이터베이스에 전달된다.

변수 설정은 지도화 시스템이 처리하기 위한 변수들을 선택하는 과정으로 클라이언트에 보내질 여러 변수들 중에서 색상합성등에 사용할 세 종류의 변수만을 선택할 수 있도록 한다.

〈그림 4-5〉는 〈그림 4-4〉의 폼에서 전달되는 정보를 이용하여 클라이언트에 보내진 애플리케이션이다. 이 중에서 상단부에 보이는 텍스트필드(textfield)들은 클라이언트에 보내진 각종 자료를 표현한다. 사용자는 마우스 포인터를 특정 공간 자료위에 올려놓음으로써 그 공간 자료를 참조하는 각종 텍스트 자료를 볼 수 있다. 예를 들어, 신림동을 표현하는 폴리곤(polygon) 위에 마우스 포인터를 올려놓는 경우에 신림동의 각종 인구 자료, 질병발생수 및 신뢰도를 표현하는 각종 수치들이 텍스트 필드에 나타나게 된다.



〈그림 4-4〉 지도화 시스템을 시작하기 위한 CGI Form

② 색상 합성을 통한 변수값의 표현

코로플레스 맵에서는 동별 변수값을 칼라 합성 방식으로 표현할 수 있다. 이는 인간이 갖는 시각적 인지 능력의 한계로 인해 미세한 칼라의 변화를 알아내기 힘들다는 문제[28]와 서로다른 여러 가지 변수를 각각의 고유 색상으로 하여 합성할 수 있다는 장점으로 인해 더욱 중요한 기능이다.

지도화 시스템은 ① 선택된 세 가지 변수를 적색(red), 녹색(green), 청색(blue)의 세 가지 색상으로 합성하여 보여줄 수 있는 기능과 ② 특정 변수만을 선택하여 표현하는 기능, 그리고 ③ 색상(hue), 채도(saturation), 명도(value)를 통해 표현하는 기능을 가지고 있다. 이를 통해, 변수값의 표현에 있어 다양한 유형을 제공한다.

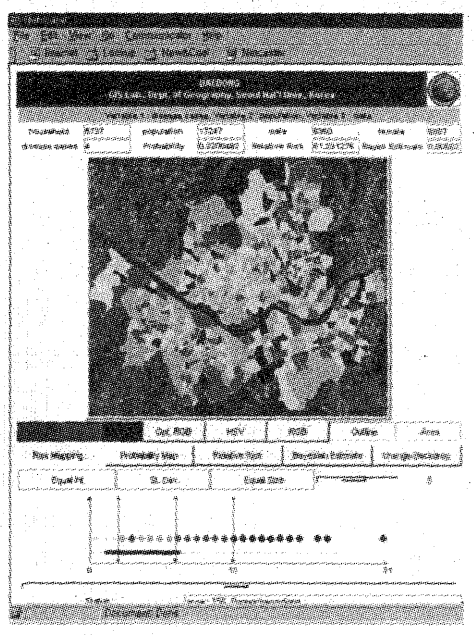
③ 자료의 분류

분류(classification)를 통해 인간, 장소, 시간 등이 유사한 형태의 틀 속에 들어감으로써 같은 집단안에 들어있는 것들끼리의 상호작용과 서로다른 집단간의 관계 등을 알아볼 수 있다. 코로플레스 맵의 분류는 면 단위로 집계된 자료를 특정 개수의

집단으로 분류한 후에 면 심볼이 갖는 색상으로 할당하는 과정을 말한다(Michael, 1997).

분류의 과정에서는 클래스(class) 수나 영역(category)을 어떻게 나눌 것인가가 주요 관심사가 된다. 일반적으로는 구 또는 동과 같은 행정단위[29]를 기준으로 하여 수집된 자료를 이용한다.

지도화 시스템에서는 여러 가지 분류 방식 중에서 동일간격 분류(equal interval), 표준편차를 이용한 분류(standard deviation), 빈도수를 이용한 분류(equal frequency)를 구현하였다.



〈그림 4-5〉 지도화 시스템에서의 표현

〈그림 4-5〉의 상단부에 보이는 서울시 지도는 표준편차를 이용한 분류 방법과 5개의 클래스 수를 이용하여 만든 코로플레스 맵이다.

그림 하단부의 1차원 산포도는 분류된 자료가 전체 데이터의 범위 중에서 어디에 해당하는지를 보여준다. 분류된 각 클래스는 색상으로 구별되며, 산포도의 클래스 경계선 및 색상은 범례의 역할을 대신할 수 있다.

상단부의 코로플레스 맵과 하단부의 산포도에 표현된 그래픽 요소들은 마우스 포인터가 움직임에 따라서 동일한 공간단위를 표현하므로 자료의 확인에 유용하게 쓰일 수 있다.

구현된 세 가지 분류방식 중 표준편차를 이용한 분류는 질병발생 자료가 정규분포를 따른다고 보기 어렵기 때문에 유용한 방법이라고 할 수 없으며 인구학적 자료의 분류 기법으로 이용되는 것이 좀 더 타당할 것으로 보인다. 반면에 구간수를 이용한 분류 및 동일간격 분류는 질병발생 자료의 전반적 분포양상을 탐색하는 데 유용하게 이용될 수 있다.

④ 위험도 수치의 지도화

질병발생 자료를 인구에 대한 비(rate or proportion)[30]로 표현했을 경우에는 그 비율의 디스플레이가 분석의 첫 번째 단계가 된다. 그러나, 단순 비의 표현은 인구값의 함수가 되므로 표현 및 해석의 오류가 생길 가능성이 높다. 따라서, 이를 해결하기 위한 여러 가지 대안 중 상대적 위험도(relative risk)[31]의 표현은 보건지리학에서 일반적으로 사용하는 방법이다. 그러나, 질병발생수가 극히 적은 경우에는 값이 과장되어 나타날 수 있다는 문제점[32]이 있으며, 따라서 자료값의 유의성을 판단하기 위해 확률지도(probability map)[33]가 사용될 수 있다. 질병발생은 각 공간단위별로 그리 자주 나타나는 현상이 아니므로 포아송 분포를 따른다는 가정하에서 실제 자료가 유의한 값인지 아닌지를 포아송 분포의 값과 비교하여 극단적 값을 가진 지점들[34]을 색상으로 표현하였다.

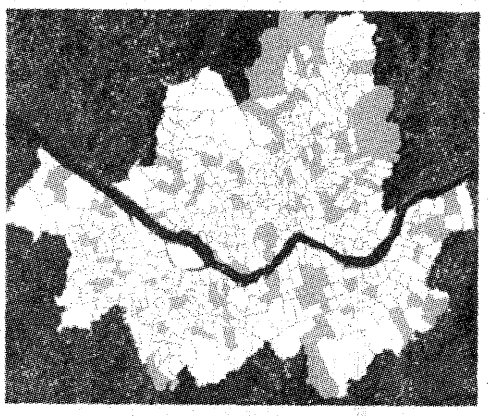
확률 지도가 갖는 한계[35]를 보완하기 위해 베이시안 측정치(baysian estimate)를 사용할 수 있다.

위험도 수치 및 그 자료의 유의성을 판단하기 위한 수치들은 통계분석 서버에서 처리하여 넘겨준 것으로 미들웨어의 주 서버에서 이를 코로플레스 맵의 형태로 표현한다.

〈그림 4-6〉은 베이시안 측정치를 지도화한 것으로 색상으로 표현된 지역은 측정치의 분포에 있어

서 90%이상이거나, 10%미만에 해당하는 수치를 가진 지역들이다.

이 지도와 위험도 수치를 보여주는 지도를 비교하는 작업 등을 통해서 신뢰도가 낮은 지역들을 제거하는 데 유용하게 이용될 수 있다.



〈그림 4-6〉 베이시안 측정치(Baysian Estimate)를 이용한 위험도 지도(Risk Map)의 신뢰성 표현

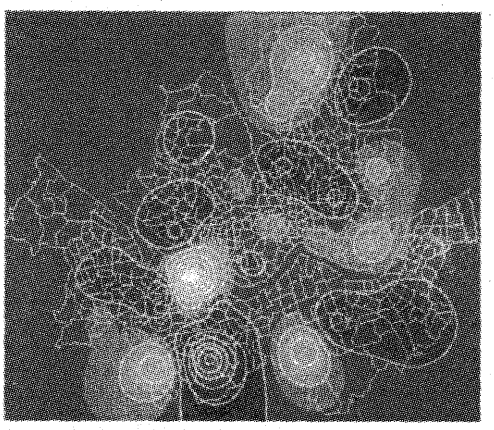
⑤ 등치선도의 제공

등치선도는 질병발생의 전반적 분포를 한눈에 파악할 수 있는 표현이다. 등치선도의 생성은 GIS 서버에서 이루어지며, 구별로 집계된 자료를 보간하고 이 자료를 이용하여 등치선도를 만든 후에 주 서버로 넘겨준다.

생성된 등치선도는 보간된 이미지와 중첩된 것으로 동경계와 중첩해서 관찰함으로써 질병이 집중되어 발생하는 지점을 확인할 수 있도록 도와준다.

〈그림 4-7〉은 배경 이미지를 위성 이미지에서 등치선도로 바꾼 것이다. 보간에는 처리의 효율을 위해 구별로 집계된 자료를 이용하였으며, 필요한 경우 동별로 집계된 자료를 보간에 사용할 수 있다. 자료값을 색상으로 표현하는 과정에서 자동스케일 변환(autoscaling)이 이루어진다.

〈그림 4-7〉에서 상대적으로 높은 질병발생율을 보이는 지역은 다양한 필드 조사 등을 통해 그 원인이 규명될 수 있을 것으로 생각된다. 즉, 본 논문에서 제시한 등치선도는 공간적 분포 경향을 보여주기 위한 단서가 될 수 있다.



〈그림 4-7〉 등치선도와 동경계의 중첩

⑥ 축척의 변화

자료의 집계는 축척에 따라 다른 양상을 보일 수 있다. 집계된 질병발생 자료의 기본 공간단위는 행정구역이므로 그 축척은 동과 구로 크게 나눌 수 있다. 공간 단위의 세밀함에 따라 전체적 분포 양상이 달라질 수 있으나, 구현된 지도화 시스템에서는 구별 자료를 집계하지 않으므로 단지 확대/축소(Zoom in/out) 기능만을 이용할 수 있도록 하였다. 확대/축소 기능은 인간의 시각 능력이 가지는 한계를 극복하기 위한 하나의 방법이라고 할 수 있으며, 특히 조밀하게 분포된 공간객체들(densely populated spatial objects)의 자료값을 탐색하는 경우에 유용하다. 현재 GIS-웹 인터페이스를 설계하는 대부분의 사이트가 지도 자료의 가시화 및 표현에 있어서 확대/축소 기능을 중요한 요소로 꼽고 있다는 사실은 이것이 자료의 탐색에 있어 갖는 중요성을 말해주는 것이라 할 수 있다(Steve, 1997).

2. 질병발생 자료집계 시스템

1) 질병발생 데이터베이스 접근

자료 분석의 초기 단계에서는 테이블이나 히스토그램과 같은 1차원 그래픽으로 기술하는 것이 일반적이며, 공간 분석으로 확장되는 과정에서 지도로 표현하는 방식이 도입된다.

그런데, 두 방식 모두 웹으로부터 질병발생 데이터베이스에 접근할 수 있어야 하므로 데이터베이스 접근방법의 선택은 웹 인터페이스 구현에 있어 중요한 요소가 된다.

본 연구에서는 클라이언트/서버 방식에 있어서 유연한 확장을 제공해 줄 수 있는 방식으로 공용 게이트웨이 방식을 사용하였다.

질병발생 데이터베이스에 저장된 의료보험자료는 원래의 의료보험자료로부터 스키마(database schema)[36]를 설계한 후에 만들어진 것이다.

2) 자료집계 시스템에서의 그래프 표현

주소를 기준으로 지역별 질병발생수를 집계하는 것은 공간성을 고려한 분석에 있어 필수적인 작업 중의 하나이므로, 사용자가 설정하는 매개변수만으로도 쉽게 이루어지도록 만들 필요가 있다. 이때, 결과물은 테이블과 텍스트 유형 중 선택할 수 있어야 하며, 이는 산출된 결과물을 다른 분석의 입력으로 재사용할 수 있도록 한다.

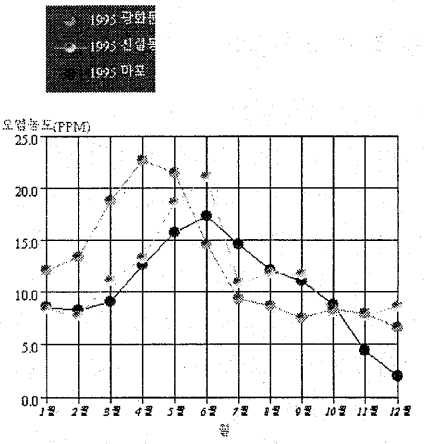
자료 분석의 초기 단계에서 일반적으로 제공할 수 있는 가시화 방안 중 하나가 그래프이다.

질병발생 자료와 질병발생의 요인을 추적할 수 있는 환경 정보들이 시간과 관련된 변수(timestamp)를 포함하고 있다는 사실은 이를 이용한 가시화 및 분석이 다양하게 이루어질 수 있음을 의미한다.

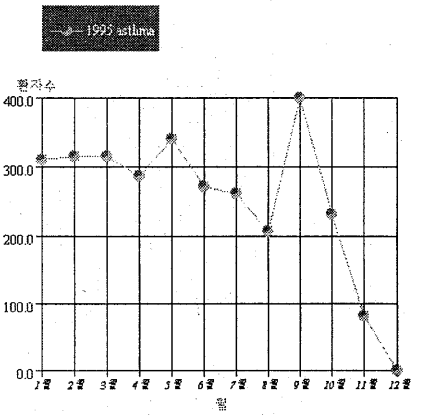
〈그림 4-8〉은 1995년 오존 평균치의 월별 변화를 보여주는 그래프이다. 그래프 시스템에서는 서울시 대기오염 농도를 측정하는 각 측정소와 연도를 선택할 수 있는 CGI폼을 제공한다.

사용자는 원하는 수의 측정소를 선택할 수 있으며, 그래프 시스템은 이 정보를 받아서 각 측정소에서의 월별 추이를 자바 애플릿으로 표현한다.

이 그래프를 통하여 공간적으로 분포되어 있는 측정소별 농도의 차이를 비교할 수 있다. 각 오염물질 별로 공간적 차이가 심한 경우와 그렇지 않은 경우가 확인될 수 있다.



〈그림 4-8〉 1995년 오존 평균치의 월별 변화



〈그림 4-9〉 1995년 천식 환자수의 월별 변화

〈그림 4-9〉에서 보이는 특정 질병발생수의 월별 변화는 전체적인 질병발생의 추이를 보여주는 데 유용하게 사용될 수 있다. 특히, 감염병 발생과 관련하여 자료의 시공간적 추이를 보여주는 전 단계에서 이용될 수 있을 것으로 보인다.

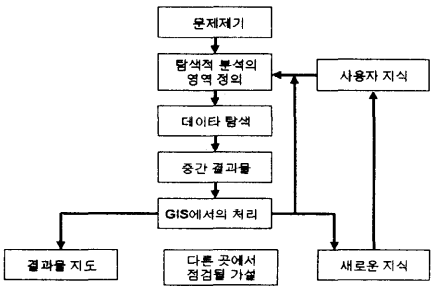
환경오염 관련 자료로 이용되는 대기오염 농도

변화와 질병발생수의 변화를 좀더 쉽게 탐색할 수 있도록 두 변수간의 상관 정도를 보여줄 수 있다. 이 그래프를 이용하여 개략적인 상관정도를 파악할 수 있다.

3. 웹-통계모듈 시스템

1) 경험적 자료 탐색

GIS에서는 그래픽 요소가 강조되므로 그 결과물이 지도와 같은 이차원 그래픽으로 만들어질 수 있다. 이때, 결과물은 단순히 통계나 모델의 인자에 관한 수치값으로 나오지 않는다. 따라서, 이러한 GIS의 특성은 가시화에 강조를 두는 또다른 공간분석 패러다임을 제공한다(Manfred et al, 1996 : p28-29). GIS에서 사용하는 경험적 자료 탐색은 반복적으로 적용되며, 탐색을 반복할 때마다 인간의 직관과 지식이 입력으로 들어가고 그 가시적 결과물을 통해 새로운 지식을 습득하게 된다.



자료원: Manfred et al, 1996 : p29.

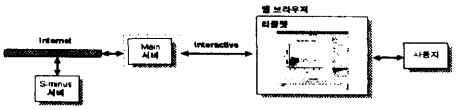
〈그림 4-11〉 경험적 자료탐색을 위한 개념적 모형

2) 웹을 통한 질병발생 자료의 탐색

일반 사용자 및 분석가가 행하고자 하는 공간 분석은 다양하고 복합적이므로 일반적인 유형의 인터페이스를 제공하는 방식만으로 사용자의 요구를 모두 만족시킬 수는 없다. 따라서, 지도화 시스템이 가지는 질병발생 자료 탐색의 한계를 극복하기 위해 통계모듈과 웹과의 인터페이스를 구현할 필요가 있다.

본 연구에서 사용된 통계모듈인 S-minus는 개발자가 공간분석 함수를 추가하기 쉽다는 장점을 가지고 있으므로 웹과의 인터페이스를 개발하였을 경우에 질병발생 자료의 다양한 공간분석에 유용하게 이용될 수 있다.

3) 대화식 웹-통계모듈 시스템



〈그림 4-12〉 웹에서의 대화식 통계분석을 위한 연결구조

대화식 환경을 제공하는 웹-통계모듈 시스템은 사용자의 명령을 서버가 받아서 통계분석 서버에 전달하는 연결구조를 이용한다. 이 구조 역시 전체 시스템 구조의 개념에서 크게 벗어나지 않는다.

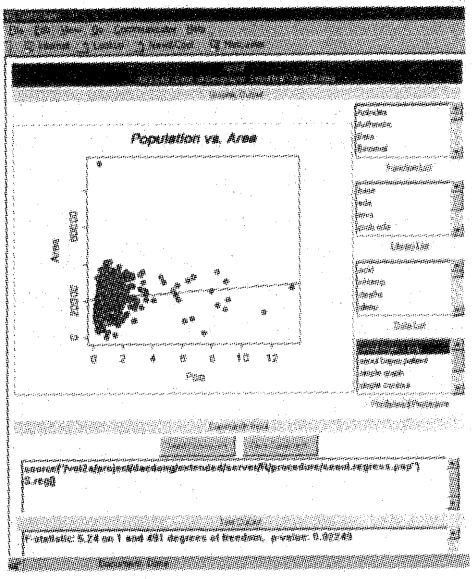
〈그림 4-13〉은 구현된 웹-통계모듈 인터페이스를 보여준다. 오른쪽의 리스트박스(listbox)에서는 통계모듈이 지원하는 함수와 라이브러리를 찾아보거나 선택할 수 있다. 구현을 위해 사용된 S-minus는 데이터를 객체로 저장하고 관리하므로 현재 이용가능한 객체의 종류 및 내용을 볼 필요가 있다. 따라서, 객체 탐색을 위한 뷰어(viewer)가 필요한 경우 추가될 수 있다. 사용자는 데이터 객체를 선택함으로써 자동으로 명령 입력이 가능하다.

사용자가 입력한 명령이 완성되면 이 명령은 서버로 보내질 수 있으며, 그 결과물은 〈그림 4-13〉에서 보는 바와 같이 그래픽이나 일반 텍스트의 유형으로 표현된다. 그림에서 표현된 것은 서울 각 동의 인구와 면적간의 회귀선(상단의 그래픽)과 텍스트 출력물(하단의 텍스트)이다.

서버쪽에서 발생할 수 있는 다양한 에러나 경고 메시지는 디버깅을 위해 결과 출력 텍스트와 함께 사용자가 볼 수 있도록 설계되었다.

이 하부시스템에서는 공간 통계를 위한 모듈을

추가함으로써 GIS 모듈의 기능 보완을 위한 확장이 가능하다.



〈그림 4-13〉 웹-통계모듈 인터페이스

Ⅴ. 결론

본 연구는 질병발생 자료를 지도화하고 웹으로 서비스하기 위한 방안을 마련하는 데 초점을 두고 있다. 이를 위해 질병발생 자료가 기초 데이터로 자주 사용되는 보건 지리학 분야에 있어 GIS를 활용할 때 필요한 개념적인 요소들을 살펴보았다. 또한 구체적인 시스템의 설계를 위해 시스템의 사용자 요구사항을 살펴보았으며 이를 만족시킬 수 있는 시스템의 설계 및 구현 방식을 모색하였다.

연구 결과를 요약하면 다음과 같다.

첫째, 원자료로부터 질병발생 정보를 만들어내기 위한 시스템 아키텍쳐를 도출하였다. 아키텍쳐의 도출에 있어 분산된 이질적 시스템들 간의 효율적 연계를 돕기 위한 방안으로 클라이언트/서버 방식을 도입하였다.

시스템 아키텍쳐를 도출함으로써 분산된 질병발생 데이터베이스, 자료의 분석 및 가시화에 사용되는 통계분석도구, 그리고 GIS 기능을 가진 모듈 또는 관련 응용프로그램에 접근하기 위한 구체적방식이 제안되었다.

둘째, 질병발생 자료를 지도화하고 웹 브라우저로 제공할 수 있는 웹 인터페이스를 설계 및 구현하였다. 티클 플러그인을 이용하여 구현된 지도화 시스템은 면 심볼을 이용하여 질병발생 자료를 지도화한다.

이 시스템은 웹 사용자가 대화식으로 처리할 수 있는 기능들을 갖추고 있다.

셋째, 질병발생과 관련된 여러 가지 공간 및 비공간 자료를 탐색할 수 있는 통합 시스템을 구현하였다.

이 시스템은 지도화 시스템의 기능을 보완하기 위해 웹-DBMS 인터페이스와 웹-통계모듈 인터페이스를 하부 시스템으로 포함한다. 웹-DBMS 인터페이스는 서울의 동을 기본 공간단위로 하여 특정 질병의 발생수를 집계하고, 그래프를 이용하여 그 변화 추이를 보여줄 수 있다. 웹-통계모듈 인터페이스는 질병발생 정보의 사용자 및 분석가가 요구하는 다양한 공간통계적 분석을 웹 상에서 대화식으로 지원할 수 있다.

본 연구에서는 단순 통계 수치에 그치던 질병발생 자료의 분석 결과물을 GIS를 이용하여 지도화하였으며, 웹 상에서 정적인 질병발생 지도를 제공하는 방식을 발전시켜 대화식으로 제공할 수 있도록 하였다.

연구의 결과로 도출된 시스템 아키텍쳐는 앞으로 우리나라에서 구축될 질병발생 정보 시스템의 분석과 배분 기능을 담당하는 하부시스템 구조로 제

시될 수 있을 것으로 생각된다.

앞으로 본 논문에서 이용한 지도화 기법 이외에 보건지리학에서 사용하는 다양한 개념 및 기법이 GIS를 이용하여 질병발생 자료에 적용될 수 있을 것으로 생각된다.

〈註〉

1) 의료지리학, 의학지리학 등의 용어가 사용될 수 있으나, 본 논문에서는 보건지리학이라는 용어를 사용하였다.
2) 의사소통 도구로서의 지도에 관한 개념의 변화는 Arthur et al. 1976 : pp. 23-42 에 잘 기술되어 있다.
3) 일반적으로는 GIF나 JPG 파일 포맷으로 제공되며, 한 번 웹 브라우저로 전송된 이미지는 사용자가 웹 상에서 조작, 변형하는 것이 불가능하다.
4) 가톨릭대학교 의과대학 예방의학교실, 1997년 5월, 전염병 및 예방접종 Surveillance 체계 개발.
5) 폐암, 천식, 심장질환 등 암 발생 자료.
6) 히포크라테스 시절부터 서양의학은 건강에 영향을 미치는 요인으로 장소를 인식하고 있었다. (Luc, 1994).
7) 마취과 의사였던 그는 큰길 주변에 위치한 공중우물을 사용하는 가정에서 높은 밀도로 콜레라가 발생한다는 것을 발견하였다. 그는 우물의 오염된 물이 원인일 거라고 생각하고 정부가 우물을 폐쇄하도록 만들었다. 이를 통해 새로운 콜레라의 발생이 급격히 줄어들었다.
8) AIDS Data Animation Project : Consortium of Earth Sciences Information Network(CIESIN), CDC(U.S. Centers for Disease Control and Prevention)참조.
9) 이 개념적 모델은 특히 전염병 발생의 기술 및 분석에 유용하게 사용될 수 있다.
10) IDRC(International Development Research Center)는 보건 연구에 대한 GIS 사용을 잘 기술하고 있다.
11) 1994년의 한 연구는 "보건지리학에서의 GIS 응용"에 관한 기존 문헌 연구에서 단 한 건의 연구 문헌만이 존재한다고 밝히고 있다(Luc, 1994).
12) 미국에서는 문제되는 질병에 대한 발생과 사망수를 질병관리소(CDC)가 각 주로부터 보고받아 분석함으로써 효율적 모니터링을 하고 있으며, 이를 매주 단위로 발간하여 관계 인사들에게 배포하고 있다. 우리 나라의 경우에는 감시체계가 미비하여 단순 통계자료조차 제대로 배포하지 못하고 있다. 전염성 질환의 경우 1990년부터 국립보건원이 발간하기 시작한 "감염병 발생정보" 정도가 전부이다(김정순 외, 1996 : p.14).
13) 여러 가지가 있으나 특히 하드웨어 인터페이스의 표준화, 공간데이타 변환표준의 개발, 범용 공간데이타베이스의 축적, 대용량 지리정보데이타를 위한 네트워크 기반시설 등의 기술적 문제가 크게 작용한다.
14) 질병관련 정보를 체계적이며 지속적으로 수집하고 분석하여 그 결과를 관련자에게 배분함으로써 궁극적으로 질병의 관리 및 예방활동에 적용하는 것을 말한다(신의철, 1997, 전염병 surveillance 및 sentinel surveillance 체계, 환경성 질환의 감시체계 구축에 관한 세미나).
15) 감염성 질환 발생 정보의 분석 목표는 전염병의 확산을 예측하는 것으로 우리나라의 경우에는 우선 감염병 발생 자료의 수집 및 분석을 통한 감염병 관리사업에의 이용과 더불어 효과적인 정보 배분에 초점을 두고 있다(감염병 발생 정보, 1996, Vol. 7, No. 7 (참조)).
16) 비감염성 질환 발생 정보의 분석은 주로 질병 발생 인자중에서 환경적 요인의 영향 정도를 파악하는 데 초점을 두고 있다.
17) 지식 정도 및 시스템 사용의 목적 등...
18) GIS 세미나, 1997, 인터그래프 코리아 (구조 참조).
19) David Flanagan, 1996, Java in a nutshell, O'Reilly & Associates. pp.4-5.
20) Ousterhout, J. K., 1994, Tcl and the Tk toolkit. Wokingham : Addison-Wesley.
21) 어떤 도구의 결과물을 외부의 소프트웨어를 이용하여 처리하는 과정을 포함한다. 이 외부 소프트웨어는 결과물을 HTML로 만들어줌으로써 웹 브라우저가 내용에 접근할 수 있도록 한다.
22) 전염병에 관한 점 패턴 분석의 예는 1854년의 John Snow 까지 거슬러 올라간다. 점 패턴의 연구에 사

용된 몇몇의 통계기술은 (Gesler, 1986)에서 확인할 수 있다.

23) 코로플레스 맵은 19세기 초에 소개되었으며 미국의 인구조사국(the Bureau of the Census)에서 통계수치의 지도화를 위해 사용되어 오던 것이 점차 지리학자와 지도학자가 가장 좋아하는 지도화 수단이 되었다. 이 이름은 그리스어의 choros(place)와 pleth(value)에서 유래된 것으로 면지도화(area mapping)라고도 불린다. (Borden, 1993 : p 125).

24) 공간 분석에 관한 논문 76개 중에서 54%에 해당하는 것이 분석 도구로 코로플레스 맵을 사용했다는 검색 결과가 있기도 하다. (Flavio et al, 1994).

25) 넓은 지역은 신뢰도가 높은 측정치를 가진 좁은 지역의 중요도를 떨어뜨린다.

26) 지도 변환(map transformation)은 인구의 크기에 따라 지역 경계를 변환하는 방법이다. 그러나, 변환된 지도는 해석하기가 어려운 경우가 많으며, 지리적 스케일이 큰 경우에는 군집성을 찾아내기가 어렵다는 단점이 있다.

27) 기본 공간 단위의 인구나 면적의 영향을 보여주기 위해서는 원 심볼을 사용하는 cartogram이 흔히 사용되지만, 이것은 공간단위의 수가 적을 때만 효과적이다. 즉, 서울시 동 526개를 사용하는 경우에는 개개 원의 수가 너무 많기 때문에 그 패턴을 쉽게 알아보기 힘들다는 단점이 있다.

28) 인간의 색상 인지 능력을 극복하기 위해서는 log scale로의 변화를 통해 값들간의 차이를 과장하는 방법이 쓰일 수 있다.

29) chorogram 이라고도 불린다.

30) Ri=Yi / Ni (단, Yi는 i동에서의 질병발생수, Ni는 i동의 인구).

31) μi(expected number of deaths in each area)= Ni * (ΣYi / ΣNi)Relative risk = yi / μi * 100
(단, Yi는 μi 라는 기대값을 만들어내는 독립적 포아송 확률 변수로부터의 관측치).

32) small number problem.

33) 부록 참조.

34) 90%이상인 값과 10% 이하인 값으로 정하였다.

35) probability map은 first order의 영향만을 측정하지만, 실제로 극단적 값들은 second order, 즉 공간 의존성(spatial dependence)의 영향을 받을 수 있다. (Trevor et al, 1995).

36) 스키마 설계는 기 정의된 것을 사용하였으므로 자세히 다루지 않는다.

〈參考文獻〉

유근배, 1996, 국가지리정보시스템의 네트워크 설계를 위한 기초연구, 국토개발연구원.

채기준, 김준년, 1995, 데이터통신 및 분산망, 희중당.

가톨릭대학교 의과대학 예방의학교실, 1995년 12월, 법정전염병 보고실태와 보고행태 관련 특성에 관한 연구.

Kenneth J. Dueker, Ric Vrana, 1995, Systems Integration : A Reason and a Means for Data Sharing, Sharing Geographic Information, The State University of New Jersey.

Manfred Fischer, Henk J. Scholten, David Unwin, 1996, Spatial Analytical Perspectives on GIS, GISDATA 4, Taylor & Francis.

Trevor C. Bailey, Anthony C. Gatrell, 1995, Interactive spatial data analysis, John wiley & sons.

Andrew D. Cliff, Peter Haggett, 1988, Atlas of Disease Distributions, Blackwell Publishers.

Richard Thomas, 1992, Geomedical Systems, Routledge, London.

Christine E. Dunn, Simon P. Kingham, 1997, Modelling air quality and the effects on health in a GIS framework, Innovations in GIS (3), Taylor & Francis.

Luc Loslier, 1994, GIS from health perspective, Proceedings of an International Workshop, Colombo, Sri Lanka.

Scholten, H.J., de Lepper, M.J.C., 1991, The benefits of the application of geographical information systems in public and environmental health, WHO Statistical Quarterly, 44(3).

Kabel, R., 1990, Predicting the next map with spatial adaptive filtering, In proceedings of the fourth international symposium in medical geography, Norwich, 16-19 July, University of East Anglia,

Norwich, UK.

Arne Koschel, Ralf Kramer, Ralf Nikolai, 1996, Third International Conference/Workshop Integrating GIS and Environmental Modeling, Santa Fe, New Mexico, USA.

Buehler K., 1994, The Open Geodata Interoperability Specification, OGIS Project Document 94-025R2.

Blaine H. Berger, 1997, The Web Enabling of Spatial Information Technologies, Proceedings of Eleventh Annual Symposium on GIS, Vancouver, Canada.

Qi Dai, Larry Evans, Mike Shank, 1997, Internet Interactive GIS and Public Empowerment, Proceedings of Eleventh Annual Symposium on GIS, Vancouver, Canada.

Hui Lin, Li Zhang, 1997, A Web-Based GIS Catalog Browser for Distributed Spatial Data Retrieving, Proceedings of Eleventh Annual Symposium on GIS, Vancouver, Canada.

Steve Morris, 1997, Web GIS and Interactive Mapping Sites, http://www.lib.berkeley.edu/UCBGIS/intergis.html Bill Thoen, 1996, WebGIS : Toy or Tool? Resource List, http://www.gisnet.com/gis/notebook/webgis.html.

Goodchild, M., Steyaert, L., Parks, B., 1996, GIS and Environmental Modelling : Progress and Research Issues, GIS World Books.

Michael N. Demers, 1997, Fundamentals of Geographic Information Systems, John Wiley & Sons.

Michael P. Peterson, 1995, Interactive and Animated Cartography, Prentice-Hall.

Hilary M. Hearnshaw & David J. Unwin, 1994, Visualization in Geographical Information Systems, John Wiley & Sons Inc.

Arthur H. Robinson, Barbara Bartz Petchenik, 1976, The nature of maps, Prentice-Hall.

Borden D. Dent, 1993, Cartography : thematic map design (3rd edition), Addison-Wesley.

Jason Dykes, 1996, Dynamic maps for spatial science : a unified approach to cartographic visualization, in innovations in GIS 3, Taylor & Francis.

Raymond S. Greenberg, Stephen R. Daniels, W. Dana Flanders, J. William Eley, John R. Boring, 1995, Medical Epidemiology (2nd Edition), Appleton & Lange.