

EMR 시스템을 위한 그리드 기술 기반의 보안성 및 데이터 접근성 향상 기법

신동민^o 신동규 신동일

세종대학교 컴퓨터공학과

gentletiger@gce.sejong.ac.kr^o{shindk, [dshin](mailto:dshin@sejong.ac.kr)}@sejong.ac.kr

A method for improving security and data accessibility in EMR systems based on GRID technologies

DongMin Shin^o DongKyoo Shin Dongil Shin

Dept. of Computer Engineering, Sejong University

요 약

지금까지 병원에서 사용하던 일반 종이차트를 벗어나 전자적으로 환자의 데이터를 기록하고 유전자 데이터를 이용하여 환자의 유사 질병까지 찾아 낼 수 있는 EMR(Electronic Medical Record 전자 의무 기록)이 개발되면서 의료계는 환자에게 더욱 신속하고 정확한 진료를 할 수 있게 되었다. 본 논문은 이에 그리드 환경을 접목하여 더 빠른 데이터 처리와 신뢰성 과 접근성을 높일 수 있는 방법을 제시한다. 첫째, 현재 기 개발된 EMR 시스템의 환경에서 인증된 사용자만이 스토리지에 접근 할 수 있도록 GSI Service를 이용하여 단일 인증 방식으로 보안성을 높이며 동시에 단 한번의 인증절차로 모든 자원을 활용 할 수 있다. 둘째, Replica Service를 이용하여 기존의 스토리지를 복제 하여 중요한 데이터 들을 보호하며 다수의 접근이 발생할 경우 처리를 분산 시킬 수 있는 방법을 제시한다. 그리드 미들웨어인 글로벌버스가 스토리지와 서버 상에서 CA인증을 담당하며 파일 전송을 담당하는 RFT는 스토리지의 Replica를 관리하는 RLS서버의 정보를 사용 하여 멀리 떨어져 있는 복제된 데이터와의 관계를 기억하고 접근시 가장 가용성이 뛰어난 머신에서 데이터를 불러온다. 이런 글로벌버스의 서비스 들은 중요하며 고용량이 데이터를 분산 시킴으로써 데이터의 지역성을 높여 재사용 혹은 동시 접근시 처리 시간을 단축 시킬 수 있다. 본 논문은 그리드 환경을 접목하여 이러한 서비스를 구현할 경우 높은 신뢰성과 접근의 신속성을 보장할 수 있다고 제시한다.

1. 서 론

최근 병원에서는 종이로 만든 기존의 차트를 대신하여 간단한 동작으로 환자의 임상기록 혹은 과거 병력까지 알아 낼 수 있는 전자차트 즉 EMR(Electronic Medical Record 전자 의무 기록)이 도입되면서 의료시설의 편리와 의료수준 개선까지 보장되게 되었다.

의료계에서는 전자적으로 기록된 환자의 임상기록을 통하여 환자에게 더욱 신속하고 정확한 진료를 할 수 있으며 이 데이터를 이용하여 데이터베이스 속의 다른 유사 질병과의 연계성 발견 등 다양한 연구 분야에 활용 되고 있다. EMR에 입력되는 데이터는 중요한 환자의 임상데이터 혹은 대 용량 유전자 데이터 이기 때문에 그만큼 높은 신뢰도와 신속, 정확성을 필요로 한다. 그렇기 때문에 EMR 을 포함한 대부분의 다른 과학기술 응용 프로그램들이 그렇듯 복잡한 연산과 대용량 데이터 처리를 위해 슈퍼컴퓨터나 클러스터와 같은 고성능 컴퓨팅 자원을 요구한다.

최근 이런 고성능 컴퓨팅 자원을 적은 비용으로 효율적으로 공급하기 위하여 그리드 컴퓨팅 기술이 도입되고 있다. 그리드는 분산되어 있는 다수의 컴퓨터에서 유휴 상태로 놓인 자원들을 이용하여 하나의 거대한 가상 컴퓨터처럼 이용 가능하도록 해주는 컴퓨팅 기술이다.

그리드를 구성하는 컴퓨팅 자원들의 관리와 이용을 위하여 그리드 미들웨어를 사용하는데, Globus Toolkit(글로벌스)은 대표적인 그리드 미들웨어이다.

본 논문에서는 기존의 EMR 시스템에 그리드 환경을 접목하여 신뢰성 확보, Replica 서비스를 이용하여 스토리지의 지역성과 데이터의 접근성을 향상 시키는 방법을 제시한다. 본 논문의 다음 구성은 EMR과 그리드 미들웨어 글로벌스에 대한 연구와, 그리드 시스템을 이용한 EMR 시스템 구조, 보안성 확보기술, 데이터 복제를 통한 접근성 향상 기술에 대한 설계를 살펴본다.

2. 관련 연구

2.1 EMR

현재 국내에서도 많은 병원들이 EMR을 널리 이용하

* 본 연구는 서울시 산학연 협력사업(JP091000)의 지원에 의해 수행되었음.

고 있다. 작게는 개원가를 중심으로 하는 협의의 EMR과 종합병원을 중심으로 하는 광의의 EMR로 구분할 수 있는데 먼저 협의의 EMR은 PC Based Program으로 주로 접수, 처방, 수납 등에 있고 작은 공간, 자원, 인력으로 가장 필요한 의무기록만을 전산화 한 것으로 할 수 있다. 종합병원의 경우 전자에 비해 광범위한 공간, 자원, 인력을 다양한 의무기록을 전산화 하여 사용하고 있다. 그 예로는 종이 차트를 수기로 작성하고 스캐닝 하여 이미지 파일로 보관하고 데이터베이스화한 검사결과와 스캐닝 한 이미지 차트를 함께 기록하는 형태이다. 또한, 종이차트 사용을 줄이고 외래기록, 처방, 수납 등 모든 진료기록을 텍스트로 입력하여 저장하는 방식이다. 후자의 경우 종이차트를 확인 하지 않더라도 환자의 지난 데이터에 쉽게 접근 할 수 있다는 장점이 있다.



그림 1 개발 단계의 BioEMR

하지만 현재 국내의 EMR 기술과 보급은 아직도 미비하다. 구체적인 표준도 제시되지 않았을 뿐만 아니라 아직까지 실무의 구현 사례가 적으며 [1] 기존에 개발 하고 있는 EMR 프로그램 역시 많은 보완을 필요로 하고 있다. 첫째 예로 들었던 프로그램은 텍스트 입력이 아니기 때문에 환자의 데이터를 활용 하는데 어려움이 있다. 또한 종이 차트를 사용하기 때문에 기존의 EMR 도입 전과 달라진 부분이 크게 없다. 둘째 예는 현재 일부 부서에서만 제한적으로 사용하고 있기 때문에 활용도가 떨어진다. 병원 전체적으로 이용이 어려운 이유는 외래 진료 중 불가피하게 종이차트를 사용할 수 밖에 없기 때문에 이것을 다시 텍스트화 하는데 어려움이 생길 수 있으며, 일부 검사 중 심전도 검사, 폐 기능 검사 등 전산에 입력이 불가능 한 부분이 있어 종이차트가 여전히 필요하다 [2]. 병원에서 사용하는 종이문서는 오랜 시간이 지나면 보관하는데 따른 유지비용과 공간적인 이유 등으로 이를 사용을

줄일 필요성이 있다. 환자 데이터를 전산화 하여 더 쉽고 빠르게 환자에게 질적으로 우수한 진료를 위해 꼭 필요한 것으로서 국내의 EMR기술도 보다 진보된 서비스 환경이 구축 되어야 한다. 따라서 본 논문에서는 EMR 기술의 확장과 빠른 서비스 제공에 대한 해결책으로 그리드 환경을 구축한다. 제시하는 기술을 이용하여 고성능 컴퓨팅 기술과 Replica 서비스를 구축하여 빠른 처리 뛰어난 보안 효과를 기대하게 할 것이라고 예상 한다.

2.2 Globus Tool Kit

글로버스는 미국ANL (Argonne National Laboratory)에 의해 개발된 그리드 미들웨어로 미국의 여러 기관들이 계산 그리드의 구축을 위해 글로버스 프로젝트에 참여 하고 있고, 현재는 그리드 표준 미들웨어로 자리잡고 있다. 글로버스는 고성능 컴퓨터와 장비, 유휴상태인 리소스를 이용하여 하나의 가상 컴퓨터처럼 사용할 수 있게 하기 위한 프로그램이다. 지역적으로 떨어져있는 많은 유휴 자원들을 초고속 네트워크를 활용해 컴퓨팅 할 수 있게 하는 장점으로 여러 연구 소를 중심으로 그리드 환경 구축을 가능하게 하고 있다 [3].

글로버스는 크게 보안, 데이터관리, 실행관리, 정보 서비스, 런타임으로 분류할 수 있으며 현재 글로버스 툴킷 4 에서는 OGSI(Open GRID Service Infrastructure)를 WS-Resource 개념으로 도입한 WSRF(Web Service Resource Framework)의 형태로 재구성 되면서 웹 서비스 스펙으로 바뀌었다. 이를 통해 OGSA(Open 그리드 Service Architecture)는 웹 서비스에 통합되고 웹 서비스 표준을 기반으로 효율적인 자원관리가 가능하게 되었다. 글로버스 툴킷 4로 발전 하면서 JAVA WS core, C WS core, Python WS core, 기존의 Globus_io를 대체해 보다 많은 개발 API를 제공하는 XIO 등, 신뢰할 수 있는 파일전송을 위한 RFT(Reliable File Transfer)API와 프로토콜을 상호 운용 할 수 있는 새로운 버전의 Grid FTP 서버 등이 기본 서비스 모듈로 추가 되었다. 아래 그림 2는 WSRF가 구현된 글로버스 툴킷 4와 OGSA, 웹 서비스 사이의 관계를 나타내고 있다 [4].

글로버스 툴킷이 사실상 그리드 표준 미들웨어가 될 수 있었던 이유는 지속적인 버전업마다 새롭고 다양한 환경을 지원하기 때문이다. 개발 당시 글로버스 툴킷은 RedHat Linux환경에서 기본적으로 지원되어왔지만 현재는 Sun Solaris, SUSE, Linux, IBM AIX, Fedora Core에서 Microsoft Windows까지 여러 운영체제에서의 호환을 지원하고 있다. 글로버스를 이용하여 그리드 환경을 구축할 경우 설치 시 요구사항이 엄격하고 전문적인 도메인이 요구되어 처음 접근하는 개발자에게 어려움이 있다.

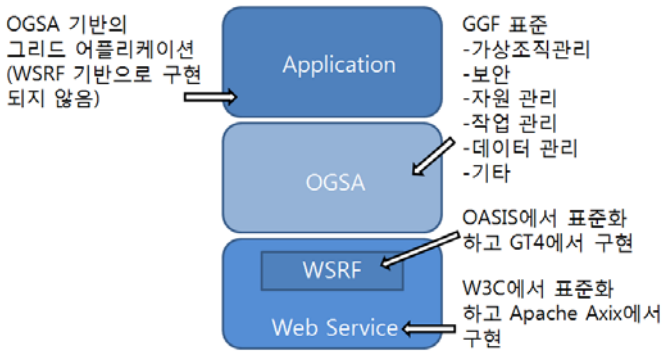


그림 2 OGSA, 글로벌스 툴킷 4, WSRF 와 웹 서비스 간의 관계

그리드 환경을 구축할 경우 주어진 양의 컴퓨터자원을 비용편의 측면에서 가장 효율적으로 사용할 수 있으며 방대한 양의 컴퓨팅 능력 없이는 풀 수 없는 문제들을 해결할 수 있는 하나의 방안이 될 수 있다. 본 논문 역시 그리드의 미들웨어로 글로벌스를 사용하여 다양한 글로벌스 서비스를 이용한 EMR 시스템의 성능을 향상시킬 수 있다고 제시한다.

3. 그리드 환경을 접목한 EMR 시스템

의료계에서 EMR시스템을 사용할 때 가장 중요한 것은 대용량 유전자 데이터에 대한 빠른 데이터 처리와 임상데이터와 같은 보안이 중요한 데이터의 보안성 예로 들 수 있다. 그렇기 때문에 EMR 시스템에 그리드 환경을 접목시킬 경우 다음과 같은 기대효과를 볼 수 있다.

기존의 EMR 시스템의 경우 의사가 환자의 데이터를 전자 차트에 기록하고 기록된 임상 데이터는 그림3과 같이 생체 데이터와 환자 임상 데이터는 각 추출기에서 정형화된 자료를 EMR 미들웨어로 전달되어 XML문서로 정형화된 데이터로 스토리지에 저장된다.

이때 기존의 시스템에서는 각각의 데이터 추출, 전송, 데이터 마이닝 관련부분에서 많은 컴퓨팅 자원을 소모하게 된다.

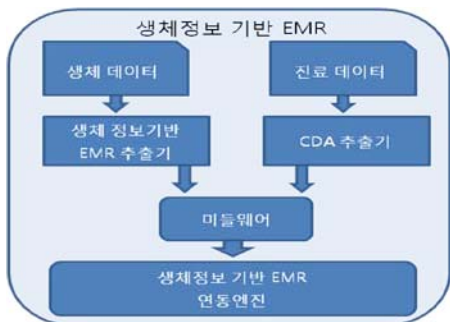


그림 3 생체기반 EMR 구조도

그림 4 는 기존에 EMR 시스템에 그리드를 접목

함으로써 사용자가 데이터 처리를 요구 할 경우 각각의 데이터간의 처리를 글로벌스가 개입하여 GRAM(Grid Resource Allocation & Management), Grid FTP등 다양한 글로벌스서비스를 이용하여 빠른 데이터 처리와 기존의 EMR 구조에서 소모되었던 자원을 다수의 글로벌스 클라이언트 머신의 유휴자원들로 대체 하여 비용적인 측면 이득을 기대 할 수 있다. 또한 기존의 EMR 시스템에 그리드를 접목할 경우 그리드를 구성하는 세 가지 요소 중 하나인 액세스 그리드를 통하여 각 클라이언트는 시간적, 공간적 제약을 받지 않고 각자의 위치에서 그룹간의 EVS, caBio, caDSR 과 같은 임상데이터, 대용량 유전자 데이터들을 공동 활용하여 협업환경을 구축 하면 멀리 떨어져 있는 다른 지역에서도 각각의 클라이언트 들은 마치 한곳에 있는 것처럼 유연한 데이터 공유가 가능하기 때문에 다양한 응용연구가 가능하다.

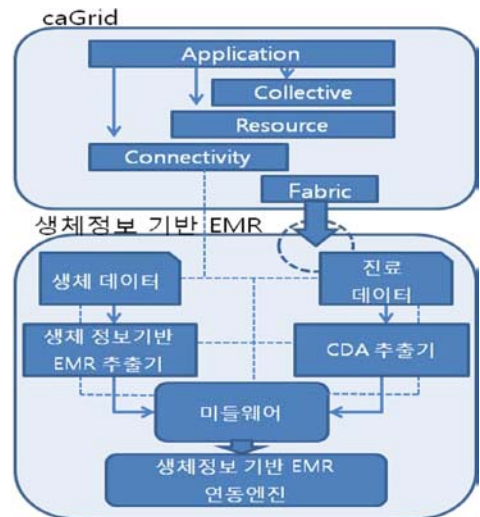


그림 4 Grid를 이용한 생체 정보 기반 EMR

4. GSI를 이용한 보안성 확보 기술

그리드 미들웨어인 글로벌스는 보안을 담당하는 부분으로 GSI (Globus Security Interface) 서비스를 제공한다. 분산된 기기종의 자원들을 공유하여 데이터를 처리하기 때문에 보안 기술은 상당히 중요한 문제가 될 수 있다. 또한 EMR 시스템의 처리되는 데이터의 특성상 환자의 인적 사항, 병력 등 보안에 많은 신경을 써야 한다. 그렇기 때문에 글로벌스의 GSI 서비스를 이용하여 EMR시스템의 보안성 확보 기술을 설계 했다.

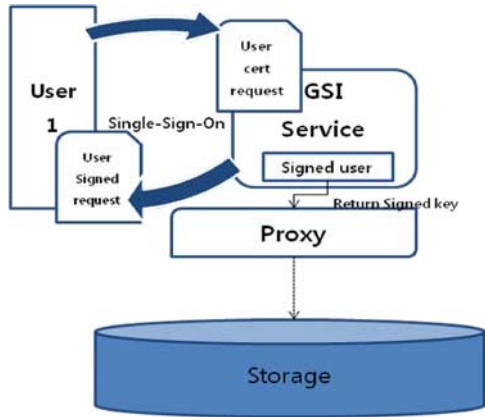


그림 5 GSI 서비스를 이용한 보안성 확보 기술

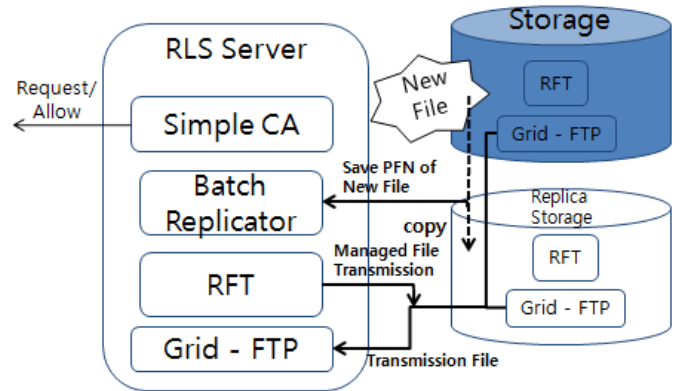


그림 6 새로운 Data의 Replica 생성 과정

먼저 글로벌스에서 GSI서비스는 사용자에게 사용의 편리성을 보장하면서 각 자원을 소유하고 관리하는 관리자의 입장에서는 자원과 데이터의 보안성을 동시에 보장한다. 이를 위해서 글로벌스는 Proxy를 이용하여 GSI는 단일 인증(Single-Signe-on) 기능을 제공한다.

스토리지에서 데이터를 요청하기 위해서는 미리 GSI 서비스를 통하여 인증되어야만 접근을 할 수 있다. GSI 서비스를 이용하기 위해서 스토리지로 사용되는 머신의 글로벌스는 Simple CA(글로벌스 인증 서비스)로 CA 인증 서비스를 실행한다. GSI의 Simple CA가 실행 되면 각각의 사용자들은 그림5와 같이 GSI의 Simple CA로 인증 Request File을 전송함으로써 인증 요청을 한다. GSI서비스를 관리하는 관리자는 인증 요청을 확인하고 다시 사용자에게 인증이 완료된 Signed Request File을 메일 혹은 Grid FTP로 전송 하는 것으로 인증이 완료 된다. 간단한 인증절차지만 이로써 해당 스토리지는 인증된 사용자만이 접근이 가능하게 된다. 이때 Proxy는 분산된 각각의 머신, 자원 혹은 스토리지에 사용자 인증을 대신 수행하기 때문에 사용자는 단 한번의 인증절차 만으로도 스토리지에 접근 할 수 있게 된다.

5. 데이터 복제를 통한 데이터 접근성 향상 기술

기존의 EMR 시스템은 하나의 스토리지에 데이터를 저장하기 때문에 다수의 사용자가 같은 데이터에 접근할 경우에도 스토리지 하나에서 모든 처리를 해야 했다.

특히 EMR 시스템 스토리지의 경우, 처리되는 데이터의 용량이 대용량이기 때문에 다수의 사용자가 데이터 처리를 요구 할 경우 많은 프로세싱 자원이 요구 된다.

글로벌스의 Replica는 동일한 파일을 그리드 환경의 다른 위치에 복제하여 생성하는 것을 말한다. Replica를 이용할 경우 데이터에 접근하는 속도를 향상시킬 수 있으며 데이터의 신뢰성을 높일 수 있다.

복제된 Replica들은 기존의 스토리지에서 예비 스토리지로 전송, 사용 하기 위하여 각각의 데이터들의 Replica 정보를 사용한다. Replica 정보는 RLS(Replica Location Service) Server에 저장된다. 그림 6은 새로운 데이터가 기존의 스토리지에 저장될 경우 자동으로 Replica를 생성해 Replica 스토리지로 저장하고 RLS에 Replica 정보를 저장하는 구조다. 사용자가 특정 데이터를 요구할 경우 RLS에 등록된 해당 데이터가 저장된 스토리지중 가장 가용성이 높은 Replica 스토리지 혹은 기존의 스토리지를 선택하여 데이터를 요청한다. RLS Server는 다수의 Replica를 관리하기 위하여 각각의 데이터에 LFN(Logical Filane)과 PFN(Physical File Name)을 사용한다. LFN은 파일에 대한 고유한 논리적 이름이고, PFN은 파일의 실질적인 위치를 URI로 나타내는 위치정보이다. 스토리지에 파일이 생성될 경우 RLS는 Replica가 저장되는 실질적인 위치를 나타내는 PFN이 추가 되고 원본 데이터와 Replica 들은 동일한 LFN을 갖음으로써 RLS로 관리 할 수 있게 된다.

그리드에서 파일의 전송을 담당하는 RFT는 Grid FTP 프로토콜을 이용하여 파일을 전송하며 파일이 전송되는 동안 파일의 전송 상태, 에러 등을 데이터베이스에 저장하여 전송되는 동안 에러 발생시 에러발생 이전의 상태로 복구하여 재전송 할 수 있는 방법을 제공한다.

그림 7은 현재 개발단계의 EMR 시스템 스토리지의 Replica를 생성하는 설계다. 설계와 같이 다수의 사용자가 같은 파일 “A”에 접근할 경우 기존의 EMR같은 경우 동일 파일에 대한 처리를 위하여 접근 지연시간이 생기게 된다. Replica로 스토리지를 확장한 경우 동일 파일 “A”에 대한 다수의 접근이 발생한 경우에도 RLS에 명시되어 있는 Replica정보를 통하여 동시 접근 사용자에게 지연시간 없이 빠른 속도로 데이터를 처리 할 수 있다.

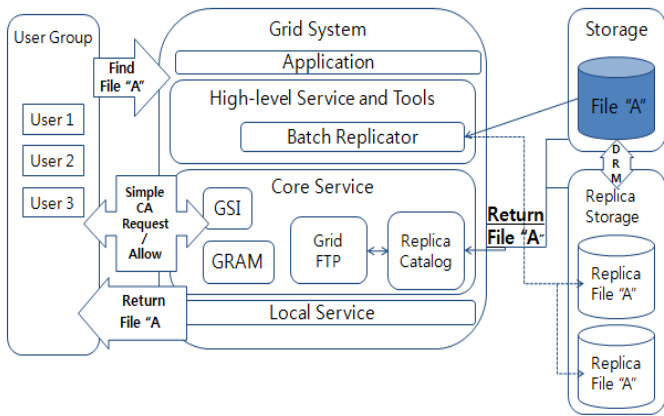


그림 7 Replica Service를 이용한 스토리지 확장

6. 결론 및 향후 연구

본 논문에서는 기존에 개발된 EMR 시스템에 그리드 환경을 접목하여 더 효과적인 컴퓨팅이 가능한 기술을 제시하고 있다.

환자들에게 보다 정확한 진료와 의료 연구에 도움을 줄 EMR 시스템이 개발 되면서 앞으로는 모든 병원에서 기존의 종이차트를 대신하여 EMR 시스템을 사용해야 할 것이다. 그러기 위해선 현재의 EMR 시스템에 대한 지속적인 보완이 이뤄져야 할 것이다.

본 논문에서는 기존의 EMR 시스템에 대한 보완점으로 컴퓨팅 능력 향상을 위하여 그리드를 접목하였고, 그리드 환경의 구축을 가정하여 GSI 서비스를 통한 보안성 향상 기술과 스토리지의 Replica를 만들어 데이터를 접근 시 처리시간을 단축시키고 더 높은 신뢰성을 보장 할 수 있는 기술을 제시했다.

향후 연구에는 본 논문에서 제시한 EMR시스템에 그리드 환경을 접목하여 GSI 서비스와 스토리지 Replica를 구현하여 기존의 EMR 시스템과의 기능성과 효율성에 대한 테스트가 필요 하며, 글로벌의 다양한 서비스들을 도입하여 더 효율적인 시스템이 될 수 있도록 EMR 시스템에 대한 연구가 계속 되어야 한다.

참고 문헌

- [1]김창환, “의료 정보화 표준화 기술 동향”, 전자정보센터, 기술동향, 3-표준화 기술, p1-16, 2008년
- [2]윤석권, 임승호, 송정영, “병원전산화를 위한 EMR 분석과 Modeling 검증”, 한국 인터넷 정보학회, 4 권, 1 호, p555-558, 2003년
- [3]위윤희, “무한 자원의 꿈, caGrid 자원 관리”, 마이크로 소프트웨어, 7 권, p230-232, 2002년
- [4]김혁호, 김양우, 이필우, “UI4GSD: 글로벌 툴킷 4 기반 caGrid 서비스 개발을 위한 사용자 인터페이스의 설계 및 구현”, 한국 인터넷 정보학회, 8 권, 5 호, p45-48, 2007년