

Original Article

한의학 연구동향 분석을 위한 페이퍼 맵 분석 시스템의 설계 및 구현

예상준¹, 장현철¹, 김철¹, 김진현², 김상균¹, 송미영³

¹한국한의학연구원 전통의학정보연구본부 정보연구센터 선임연구원,

²한국한의학연구원 전통의학정보연구본부 정보연구센터 연구원,

³한국한의학연구원 전통의학정보연구본부 정보연구센터 센터장

Design and Implementation of Paper Map for Traditional Korean Medicine

Sang-Jun Yea¹, Hyun-Chul Jang¹, Chul Kim¹, Jin-Hyun Kim²,
Sang-Kyun Kim¹, Mi-Young Song³

¹KIOM, Division of TKM Information Research, Information Research Center, Senior Researcher

²KIOM, Division of TKM Information Research, Information Research Center, Researcher

³KIOM, Division of TKM Information Research, Information Research Center, Director

Objectives: Because of the characteristics of traditional Korean medicine, there are few overseas patents and SCI papers. Researchers are struggling to analyze the research trend using existing patent analysis system and paper analysis system. We aim to build up a web-based paper map analysis system for Traditional Korean Medicine to improve the situation.

Methods: We studied the following three sub topics. First, we surveyed the research trend analysis method for traditional Korean medicine. Second, we designed a paper map analysis model and made the algorithm for it. Finally, the development of the paper map analysis system was conducted.

Results: We developed a web-based paper map analysis system which has the three main functions being analysis chart, topology map and contour map. For the rich internet application, we used the flex development tool and java integrated development environment.

Conclusion: We will provide our paper map analysis system in the OASIS through the minor changes to provide easy use. We hope that this system will be a useful tool to analyze the research trends for traditional Korean medicine.

Key Words : Paper map, research trend, traditional Korean medicine, analysis system

서론

21세기 과학기술의 발전 속도는 과거보다 무척 빠르고 변화가 심하다. 그래서 연구자가 연구개발에

대한 기획을 하거나 정부가 연구 사업에 대한 정책을 수립하기 위해서 사전조사에 많은 노력을 기울이게 된다. 연구자는 연구 분야 선정과 체계적인 연구 수행을 위해서 특허 맵과 논문 맵 분석 시스템을 이

• Received : 25 June 2010

• Revised : 9 September 2010

• Accepted : 9 September 2010

• Correspondence to : 송미영(Mi-Young Song)

대전광역시 유성구 엑스포로 483번지 한국한의학연구원 정보연구센터

Tel : +82-42-868-9454, Fax : +82-42-861-9421, E-mail : smyoung@kiom.re.kr

용하고 있으며, 기획자는 연구중복 방지와 연구의 효율성 제고를 위해서 이러한 분석방법을 이용하고 있다¹⁾. 특히 맵이란 특허정보의 서지사항을 가공하여 그 데이터를 가지고 여러 정보를 한눈에 알기 쉽게 볼 수 있도록 도표화한 것이며 논문 맵이란 논문의 서지정보를 분석하여 해당분야의 연구동향 정보를 분석하기 쉽게 작성한 것이다. 연구동향 분석에 사용되는 페이퍼 맵 분석 시스템은 대부분 외국 제품이고 Science Citation Index(SCI)에 수록된 논문에 대해서 분석기능을 제공하고 있으나 한의학계의 특성상 SCI 논문이 차지하는 비중이 타 분야에 비해서 적어서 상용 페이퍼 맵 분석 시스템을 이용하기 어려운 상황이다.

그리고 한국한의학연구원에서는 2005년 9월과 11월에 한의학계의 연구정보 공유 및 활성화를 위해서 한의과대학, 부설연구소, 대한한의학회 및 분과학회와 학술교류 협정을 체결하고 학술지를 체계적으로 수집하여 한의학 학술논문 통합검색 시스템(OASIS, 오아시스)을 개발하여 서비스하고 있다²⁾. 현재 오아시스 서비스는 48종의 한의학술 논문지에 수록된 1963년부터 2009년에 이르는 19,679건의 한의학 관련 논문을 제공하고 있다. 이러한 한국한의학연구원의 오아시스는 국내 최다의 한의학 논문을 서비스하고 있으며 한의학 연구자들이 가장 많이 이용하고 있어 한의학 페이퍼 맵 분석 시스템을 서비스하기 위한 최적의 조건을 제공하고 있다. 이러한 한국한의학연구원의 정보 인프라를 활용하여 연구동향 분석이 어려운 한의학계 상황을 개선하기 위해서 페이퍼 맵 분석 시스템 개발 연구를 수행하였다.

선행연구

1. 선행 논문

한국과학기술정보연구원의 이호신 등은 기술문헌 데이터베이스를 이용하여 기술동향을 분석할 수 있는 시스템을 개발하여 기술정보의 검색결과를 온라인상에서 시스템적으로 신속하게 분석함으로써 연구개발 및 기술개발 방향설정에 기여하여 정보수집

및 분석에 소요되는 노력과 시간을 단축시킬 수 있는 정보 분석 시스템을 개발하였다³⁾.

한국한의학연구원의 장현철 등은 전통의학 분야 연구의 중복 수행을 방지하고 체계적인 연구를 수행할 수 있도록 하기 위해 한의학 분야 프로젝트 맵 분석 시스템을 설계하고 구현하였다. 이 연구를 통해서 수집 및 분류된 프로젝트 데이터를 분석하는 시스템을 구축하기 위해서 프로젝트 데이터를 효율적으로 분석하기 위한 통계 및 분석 알고리즘을 개발하고 데이터베이스 스키마를 설계하였으며 분석 결과를 그래프를 이용하여 효율적으로 가시화하는 시스템을 개발하였다⁴⁾.

한국한의학연구원의 예상준 등은 한의학 연구동향 분석시스템 구현을 위한 모형개발을 위한 연구에서 선행연구와 선행시스템을 파악하여 한의학 연구동향 분석에 적용 가능한 분석모델을 도출하여 연구동향을 체계적으로 분석할 수 있는 기본분석, 상세분석, 복합분석의 14개 모델을 설계하였다. 이 연구를 통해서 개발된 분석모델은 한의학 페이퍼 맵 분석 시스템을 위한 서비스 시나리오 개발 및 시스템 구현에 이용되었다⁵⁾.

2. 선행 시스템

한의학 페이퍼 맵 분석 시스템을 구현하기 위한 사전 조사로써 국내외의 특허 맵 분석 시스템과 페이퍼 맵 분석 시스템을 분석하였다. Aureka, Delphion, TDA 등을 개발하는 Thomson Scientific 이 국외의 대표적인 기업이고 국내에는 특허청의 PIAS, (주)WIPS의 ThinKlear, 한국과학기술정보연구원의 KnowledgeMatrix가 대표적인 분석 시스템으로 조사되었다. 그리고 한국한의학연구원에서 2009년에 개발한 한의학 관련 연구 과제를 분석하는 Project Map Analysis System이 서비스되고 있다.

조사대상 시스템에 대해서 연산자 기반검색, 히스토리, 자연어 검색기능 지원여부와 지원하는 차트의 종류와 클러스터링, 분류체계, 리포팅 등의 부가적인 기능에 대한 분석을 실시하고 각 시스템의 장단점을 비교하여 (Table 1)과 같은 결과를 확인하였다.

Table 1. The Function Analysis of the Systems in Domestic and Foreign

		Aureka ⁶⁾	Delphion ⁷⁾	ThinKlear ⁸⁾	PIAS ⁹⁾	KnowledgeMatrix ¹⁰⁾	ProjectMap ^{2,11)}
제공기관		Thomson Scientific	Thomson Scientific	WIPS	특허청	KISTI	KIOM
개요		특허정보 관리 및 분석	특허정보에 검색 및 분석	특허맵 분석	특허정보 분석	데이터마이닝	한의학 연구과제 분석
유료/무료		유료	유료	무료	무료	무료	무료
검색	연산자 기반검색	O	O	O	O		O
	히스토리	O	O				
	자연어						
차트	Bar	O	O	O	O	O	O
	Pie	O	O	O	O	O	O
	Matrix	O	O	O	O	O	O
	3D차트		O	O	O	O	
	Contour Map	O					O
	Citation Map	O	O			O	O
클러스터링	O	O			O		
분류체계	O	O	O	O	O	O	
임포트/익스포트	O	O	O	O	O	O	
리포팅	O	O	O	O	O	O	

국의 시스템은 대부분 유료로 기능을 제공하고 있으며 기본적인 검색과 차트 기능이 충실하게 구현되어 있으며 클러스터링, 분류체계 등의 부가적인 기능도 비교적 많이 제공하고 있는 것으로 파악되며 국내 시스템은 대부분 무료로 기능을 제공하고 있으며 기본적인 기능은 충실하게 구현되어 있으나 부가적인 기능은 제한점이 발견되어 아직은 시스템의 완성도가 국외 시스템에 비해서 부족한 것으로 파악되었다.

시스템 설계

한의학 논문 분석 시스템은 차트와 테이블을 이용하여 결과를 보여주는 분석차트 기능, 그래프를 이용하여 논문의 상관관계를 탐색하는 위상맵 기능 그리고 논문의 상관관계를 한눈에 파악할 수 있게 분석정보를 제공하는 등고선맵의 세 가지 기능으로 구분된다. 모든 기능을 사용하기 전에 분석대상 논

문을 검색하는 과정이 필요하며 논문검색 기능은 오아시스 포털의 논문검색 기능과 유사하게 구현하여 사용자가 동일한 시스템을 사용하는 것처럼 별도의 기능숙지 없이 사용할 수 있도록 하였다.

1. 데이터베이스 분석

오아시스는 한의학 관련 48종의 학술지에 수록된 19,000여 편의 논문을 서비스하고 있으며 데이터베이스는 연구 분야 정보를 저장하는 TB_CATE_CODE 테이블, 학회 코드 정보를 저장하는 TB_ORG_CODE 테이블, 논문의 서지 정보를 저장하는 TB_THESIS 테이블로 구성되며 각 테이블의 연관 관계는 (Fig. 1)과 같다.

TB_THESIS 테이블이 논문에 대한 메타데이터를 저장하고 있으므로 본 연구에서 구현하고자 하는 페이퍼 맵 분석 시스템의 주된 분석대상이 되며, 선행 연구를 통해서 38개의 속성 중에서 CATE_CODE,

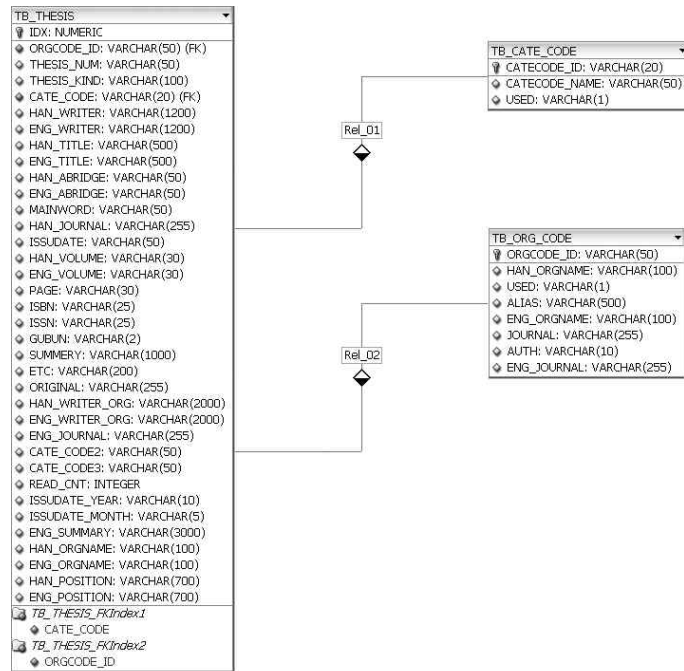


Fig. 1. The Schema of OASIS' Database

HAN_WRITER, HAN_KEYWORD, HAN_JOURNAL, ISSUDATE_YEAR, HAN_POSITION의 6개를 분석대상으로 선정하였다⁵⁾. 분석정보는 논문의 연구분야, 저자, 주제어, 학회지, 발행일, 소속기관으로 논문에 대해서 정량적으로 분석할 수 있는 속성을 대상으로 하였으며 PAGE와 READ_CNT 필드는 논문에 대한 분석정보를 제공하지 않기 때문에 분석대상에서 제외하였다.

2. 알고리즘 설계

1) 분석차트

오아시스 데이터베이스 스키마 분석결과와 선행연구⁵⁾에서 제시된 분석모델을 이용하여 한의학 페이퍼 맵 분석 시스템의 분석차트 기능을 설계하였다. 분석기능은 기본분석, 상세분석, 복합분석으로 구분되며 x축과 y축에는 분석하고자 하는 대상과 해당하는 논문수와 저자수를 랭크(값), 누계, 백분율의 형태로 생성하여 분석한다. 분석결과는 데이터유

형별로 line/column/area/bar/pie/plot/radar 차트 및 table 형태로 제공하며 상세분석은 다중 차트를 이용하여 분석결과를 제공한다.

첫째, 기본분석은 연도, 분야, 저자, 기관, 저널, 키워드의 속성별로 논문에 대한 분석결과를 생성하여 단순 차트를 이용해서 사용자에게 제공한다. (Fig. 2)의 알고리즘에 의해서 분석내용을 생성하고 각 차트는 TB_THESIS 테이블의 속성을 x축으로 하고 해당 속성을 가지는 논문수를 y축으로 구성한다.

둘째, 상세분석은 분석항목에 대한 분야, 저자, 기관, 저널, 키워드의 속성별로 상세한 분석결과를 차트를 이용하여 사용자에게 제공하며 각 차트는 분석항목을 x1축 TB_THESIS 테이블의 속성을 x2축으로 하고 해당 속성을 가지는 논문수를 y축으로 구성하는 3차원 데이터를 구성하여 (Fig. 3)과 같은 알고리즘에 의해서 분석한다.

셋째, 복합분석은 저자 상관분석과 기관 상관분석 결과를 제공하며, 저자(기관) 상관맵은 해당 저자(기

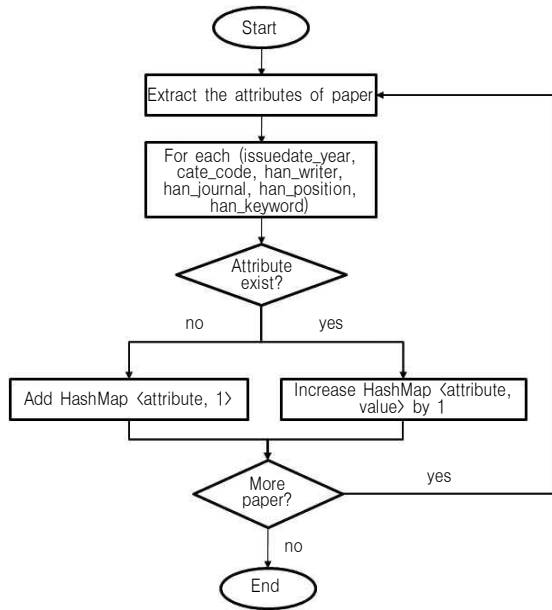


Fig. 2. The Outline Analysis and Current Status Analysis Algorithm

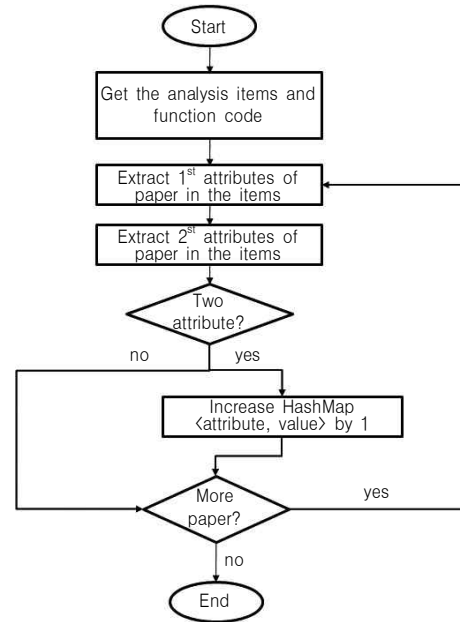


Fig. 3. The Detailed Analysis Algorithm

관들 간의 공동연구 현황에 대한 분석결과를 매트릭스를 이용하여 사용자에게 제공하며 매트릭스는 row, column을 저자(기관)로 구성하여 공동연구 현황을 분석한다.

2) 위상맵

오아시스 데이터베이스 스키마 분석결과를 적용하여 분석차트 기능에서 적용한 것과 동일한 연도, 분야, 저자, 기관, 저널, 키워드의 6개의 속성을 이용하여 위상 맵(Topology Map)을 설계하였다. 위상 맵은 (Table 2)과 같이 4단계로 구성되며 위상 맵의 특성을 효율적으로 표현하기 위해서 노드와 에지로 이루어지는 그래프를 이용하여 표시하며 ROOT (dummy node)를 루트 노드로 설정하는 경우와 해당논문을 루트 노드로 설정하는 2가지 경우의 그래프로 구분하였다.

3) 등고선 맵

논문의 등고선 맵을 작성하고 연구 활성화 정도를

분석하기 위해서 correspondence analysis와 altitude analysis를 이용하는 (Fig. 4)의 알고리즘을 이용하였

Table 2. Topology Map Graph Definition

단계	그래프 1	그래프 2
1단계	ROOT	논문
2단계	연도 분야 저자 기관 저널 키워드	논문의 해당연도 논문의 해당분야 논문의 해당저자 논문의 해당기관 논문의 해당저널 논문의 해당키워드
3단계	선택연도 선택분야 선택저자 선택기관 선택저널 선택키워드	해당연도의 논문 해당분야의 논문 해당저자의 논문 해당기관의 논문 해당저널의 논문 해당키워드의 논문
4단계	해당연도의 논문 해당분야의 논문 해당저자의 논문 해당기관의 논문 해당저널의 논문 해당키워드의 논문	

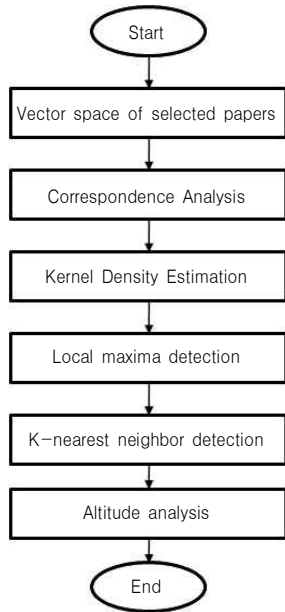


Fig. 4. Contour Map Algorithm

다. 먼저, 검색된 논문에 대한 keyword vector space 를 행렬을 이용하여 구성하고 correspondence analysis 를 적용하여 논문 간의 상관관계와 논문과 키워드 간의 상관관계를 파악한다. 도출된 2차원 좌표에 대해서 kernel density estimation function을 적용하여 3차원 모집단을 추정하여 등고선 맵을 작성하고 local maxima detection과 knn algorithm을 적용하여 해당 키워드의 위치를 계산한다. 마지막으로 등고선의 높이를 분석하여 도입, 성장, 성숙의 3단계로 구성되는 상대적 연구 활성화를 분석한다.

시스템 구현

페이퍼 맵 분석 시스템은 Flex 3.0, Java 1.6, Oracle 9i를 이용하여 개발하였으며 개발 및 웹서비스를 위해서 Tomcat 5.5를 이용하고 클라이언트 응용 프로그램 개발을 위해서 Flex Builder 3.0과 서버

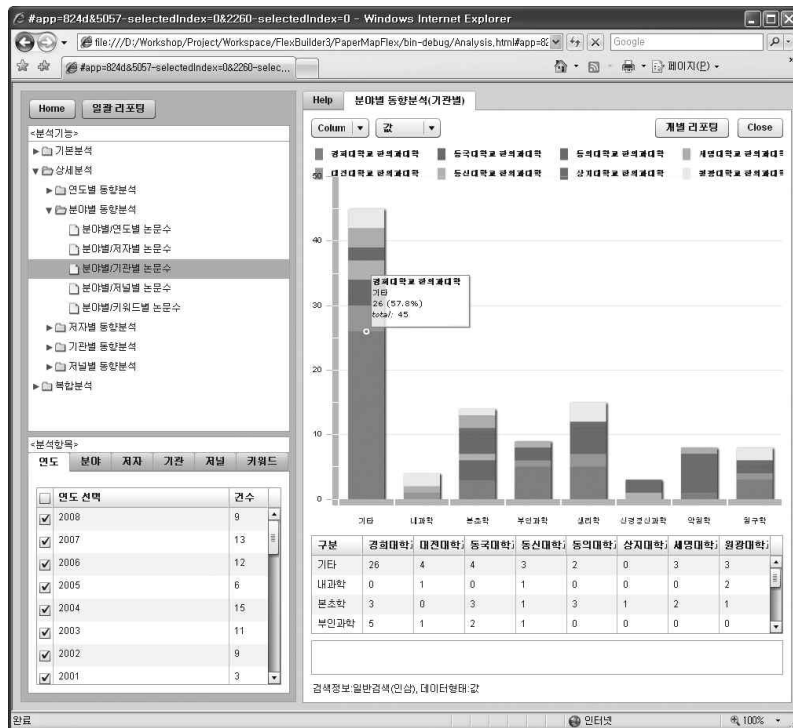


Fig. 5. Analysis Chart Screen

응용 프로그램 개발을 위해서 Eclipse Builder 3.3을 사용하였다. Flex Builder의 주요 프로그래밍 언어는 ActionScript 3.0과 MXML이고 Eclipse Builder 3.3의 프로그래밍 언어는 JAVA와 JSP이다.

첫째, 분석차트 기능은 분석기능 트리에서 특정 아이템을 클릭하면 페이지의 오른쪽에 (Fig. 5)와 같은 분석차트와 테이블로 구성된 분석결과가 생성되며 상단에는 차트모양과 데이터 형태를 선택할 수 있는 Combo Box가 있고 하단에는 분석의견을 작성할 수 있는 Text Area와 검색결과 정보를 표시하는 Text가 존재한다. 그리고 분석차트와 테이블의 데이터는 서로 연동되어 있어 테이블의 정렬순서를 변경하면 분석차트가 변경되어 표시된다.

둘째, 위상 맵의 분류리스트는 연도, 분야, 저자, 기관, 저널, 키워드의 6가지 항목으로 구성되어 있으며 오른쪽 위상 그래프의 노드 구성을 (Fig. 6)처럼 변경할 수 있다. 위상 그래프의 루트가 특정 논

문이고 분류리스트 변경된 내용을 반영하기 위해서는 적용 버튼을 클릭하고 초기화 버튼을 클릭하면 루트가 ROOT로 변경되면서 분류리스트 내용이 반영된다. 그리고 위상 맵을 탐색하기 위해서 연도, 분야, 저자, 기관, 저널, 키워드에 해당하는 노드를 클릭하면 팝업 창이 생성되어 해당 항목에 대한 정보와 논문 건수를 보여주며 사용자가 원하는 내용을 선택하면 해당 내용과 논문건수가 각 노드에 표시된다.

셋째, 등고선 맵 분석화면은 왼쪽에 등고선 맵을 분석한 내용과 오른쪽에 등고선 맵으로 구성되는 (Fig. 7)과 같이 구성된다. 등고선 맵에서 초록색 부분은 연구 활성도가 낮음을 의미하고 붉은색 부분은 연구가 활발히 수행되고 있음을 나타낸다. 그리고 페이지의 왼쪽 상단의 검색결과에는 등고선 맵을 도입분야, 성장분야, 성숙분야로 구분하고 분석한 결과의 논문건수를 표시하고 하단의 각 해당분야에는 관련된 주제가 표시되어 사용자의 등고선 맵 분석

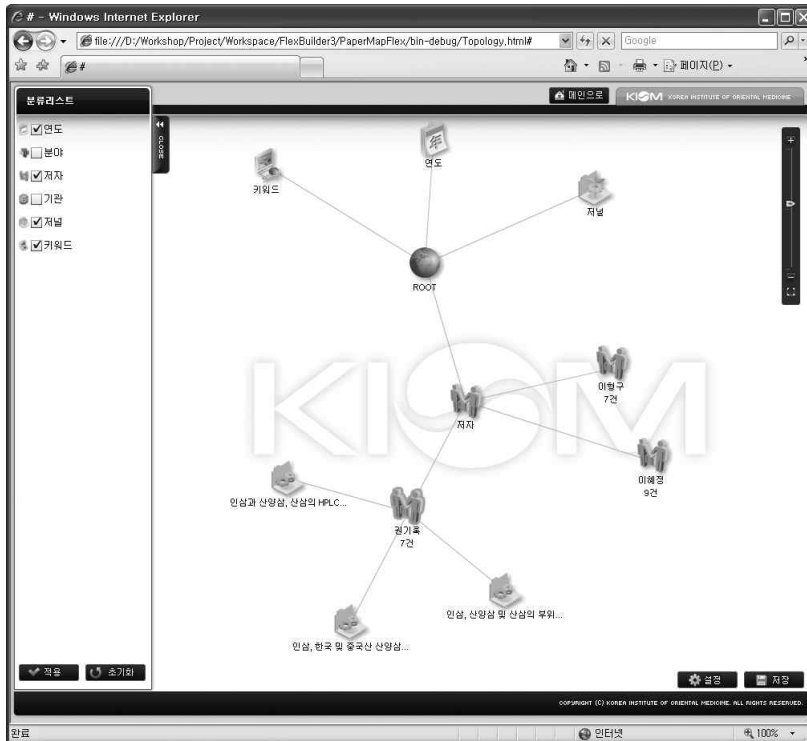


Fig. 6. Topology Map Screen

제공하도록 하였다. 또한, 웹 기반의 분석 시스템으로 구현하여 쉽고 편리하게 접근이 가능하도록 구현하였다.

향후에는 오아시스 데이터베이스에서 구축중인 논문의 참고문헌 데이터를 분석할 수 있는 효율적인 알고리즘의 개발과 논문의 상관관계를 한눈에 파악할 수 있는 사용자 인터페이스를 구현하여 페이퍼 맵 분석 시스템의 기능을 확장할 수 있으며 기존의 분석 알고리즘을 병합 및 세분화 하여 복잡적이고 세분화된 분석결과를 도출할 수 있는 방향으로 페이퍼 맵 분석 시스템의 기능을 개선할 수 있을 것이다.

참고문헌

1. Yea SJ, Han JM, Lee HJ, Song MY. Service Scenario for Korea Medicine Paper Map System Adapted from Patent Map System. Korea Computer Congress. 2009;36(1):59-63.
2. Oriental Advanced Searching and Integrated System. Available from:<http://oasis.kiom.re.kr>
3. Lee HS, Lee IH, Kim KH. Development of Technology Information Analysis System for Analyzing Technology Trend from Technology Data. The Journal of Korea Contents. 2002;2(4): 67-75.
4. Jang HC, Yea SJ, Jeon BU, Song MY. Design and Implementation of R&D Projects Analysis System in Oriental Medicine Field. Korea Computer Congress. 2009;36(1):117-122.
5. Yea SJ, Jang HC, Kim JH, Kim C, Kim SK, Song MY. Design of Models for the Korean Traditional Medicine Research Trend Analysis System. The Journal of Korea Contents. 2009;9(12):710-717.
6. Science - Thomson Reuters. Available from:<http://scientific.thomson.com>
7. Delphion Research intellectual property network. Available from:<http://www.delphion.com>
8. WIPS patent online searching system. Available from:<http://search.wips.co.kr>
9. Korean Intellectual Property Office. Available from:<http://www.kipo.go.kr>
10. Korea Institute of Science and Technology Information. Available from:<http://www.kisti.re.kr>
11. Kim SK, Jang HC, Kim JH, Song MY. Collecting Method and Implementation of Search Management System for R&D Projects in Korean Medicine Field. The Journal of Korean Oriental Medicine 2009;30(4):47-58.