

R을 활용한 정보교육관련 논문 분석

박선주

광주교육대학교 컴퓨터교육과

요 약

최근 빅데이터 분석과 함께 사회연결망에 대한 관심이 증대되고 있다. 이러한 사회연결망분석을 이용한 연구가 사회과학 영역뿐 아니라 자연과학 영역 등 여러 분야에서 다양하게 이루어지고 있다. 이에 본 논문에서는 정보교육 관련 석·박사 학위논문을 수집하여 텍스트분석과 사회연결망분석을 실시하였다. 그 결과, 모든 기간에서 출현빈도수가 높게 나오거나 지속적으로 나오는 단어가 있었으며, 기간별로 출현빈도가 갑자기 높아진 단어들도 있었다. 또한, 출현빈도수가 큰 단어가 대체적으로 매개중심성도 컸으며, 기간별 연구흐름의 특징이 있음도 알 수 있었다. 그러므로 IT 기술발전과 초·중·고등학교 정보교육과정 변화에 민감하게 정보교육 석·박사학위 논문 주제가 변화되었음을 알 수 있었다. 앞으로 기간4에서 출현빈도가 높아진 스마트, 모바일, 스마트폰, SNS, 어플리케이션, 스토리텔링, 다문화, STEAM과 관련된 연구가 지속될 것으로 예측하며, 로봇, 프로그래밍, 코딩, 알고리즘, 창의성, 상호작용, 개인정보보호와 관련된 주제도 꾸준히 연구될 것으로 예측된다.

키워드 : 빅데이터, 텍스트분석, 사회연결망분석, R, 정보교육

Analysis of Information Education Related Theses Using R Program

SunJu Park

Dept. of Computer Science Education, Gwangju National University of Education

ABSTRACT

Lately, academic interests in big data analysis and social network has been prominently raised. Various academic fields are involved in this social network based research trend, which is, social network has been actively used as the research topic in social science field as well as in natural science field. Accordingly, this paper focuses on the text analysis and the following social network analysis with the Master's and Doctor's dissertations. The result indicates that certain words had a high frequency throughout the entire period and some words had fluctuating frequencies in different period. In detail, the words with a high frequency had a higher betweenness centrality and each period seems to have a distinctive research flow. Therefore, it was found that the subjects of the Master's and Doctor's dissertations were changed sensitively to the development of IT technology and changes in information curriculum of elementary, middle and high school. It is predicted that researches related to smart, mobile, smartphone, SNS, application, storytelling, multicultural, and STEAM, which had an increased frequency in period 4, would be continuously conducted. Moreover, the topics of robots, programming, coding, algorithms, creativity, interaction, and privacy will also be studied steadily.

Keywords : Big data, Text analysis, Social network analysis, R. Information education

논문투고 : 2016-12-00

논문심사 : 2017-00-00

심사완료 : 2017-00-00

1. 서론

최근 빅데이터 분석과 함께 사회연결망에 대한 관심이 증대되고 있다. 사회연결망분석은 구조나 연결망 형태의 특징을 도출하고 관계성으로 체계의 특성을 설명하거나 체계를 구성하는 단위의 행위를 설명한다[3,15]. 트위터를 통해 본 대권주자들의 소통유형, SNS를 활용한 지방자치단체 커뮤니케이션 방안에 대한 연구[15], 사회연결망 분석을 활용한 대구의 관광지 이미지 분석[20], 소셜미디어 빅데이터로 분석한 항공사 이미지 변화[5] 등이 모두 사회연결망 분석을 활용한 연구이다.

이러한 사회연결망분석을 이용한 연구가 사회과학 영역뿐 아니라 자연과학 영역등 여러 분야에서 다양하게 이루어지고 있다. 특정분야 연구흐름을 분석하기 위해 사회연결망분석을 활용한 연구는 황동열외 (2016)는 무용연구의 빅데이터를 활용한 사회연결망 분석[6], 배성훈 외(2016)는 탄소나노소재기술 연구동향에 대한 사회연결망 분석[2], 이주연 외(2016)는 컴퓨터 분야 논문지에서 저자의 관심분야에 대한 소셜네트워크 분석[16], 배규용 외(2015)는 기후변화관련 식품분야 논문 초록 분석[1] 등의 연구들이 시작되고 있다[19].

그러므로 본 논문에서는 정보교육 관련 석·박사학위 논문의 연구흐름을 분석하기 위해 텍스트마이닝, 사회연결망분석을 실시하고자 한다. 이를 위해 학술연구정보서비스에서 이클립스를 통해 데이터를 수집한 후 R 프로그램을 활용하여 데이터를 정제하여 단어의 동시 출현빈도수 분석, 사회연결망 분석을 통해 기간별 연구주제의 변화와 단어들의 관계를 분석하고자 한다.

2. 사회 연결망 분석

네트워크는 노드와 링크로 구성되며, 사회 연결망 분석은 개체간의 관계를 기술하고 구조화하여 이들 간의 관계를 분석하는데 사용되고 있다[3,6].

사회연결망 분석에서 연결망 구조를 파악하기 위한 지표로 밀도분석(density analysis), 중심성 분석(centrality analysis), 집중도 분석(centralization analysis) 등이 있다. 밀도 분석은 연결망 노드들 사이의 연결된 정도를 분석하는 지표로 연결망 내 전체구성원이 서

로 간에 얼마나 많은 관계를 맺고 있는가를 표현하기 위한 것으로 가능한 총 관계에서 실제로 맺어진 관계의 수의 비율을 의미한다. 1에 가까운 밀도 값을 가질수록 연결망에서 연결 관계의 응집성, 결속도, 복잡성이 높다고 할 수 있다[2,20].

집중도 분석은 한 연결망 전체가 한 노드의 중심으로 집중되는 정도를 표현하는 지표이다[3,20].

중심성 분석은 각 노드가 중심에 근접한 정도를 평가하는 지표로 사회연결망 분석에서 가장 많이 사용되는 지표이다. 연결정도(degree) 중심성, 근접(closeness) 중심성, 매개(betweenness) 중심성 등이 있다. 연결정도 중심성은 연결 빈도수를 통해 한 노드에 직접적으로 연결된 정도를 측정하고, 근접 중심성은 한 노드와 직간접적으로 연결된 모든 노드간의 거리를 합산해 최단거리의 정도를 통해 중심성을 측정한다. 매개 중심성은 노드와 노드간의 중개 역할을 하는 정도를 파악해 중심성을 측정한다[7].

근접중심성 $C(i)$, 매개중심성 $C_b(v)$ 은 다음과 같이 정의된다[3].

$$C(i) = \sum_{j \neq i} \frac{1/d(i,j)}{n-1}, \quad i = 1, 2, \dots, n$$

$d(i,j)$ 는 노드 i 에서 j 에 이르는 '거리' 즉 비근접도이므로 $1/d(i,j)$ 는 노드 i 에서 j 에 이르는 근접도의 개념이다.

$$C_B(v) = \sum_{i \neq v} \sum_{j \neq v, i} \frac{g_{ivj}}{g_{ij}}, \quad v = 1, 2, \dots, n$$

g_{ij} 는 노드 i 에서 $j(\neq i)$ 로 가는 최단경로의 수이고, g_{ivj} 는 노드 i 에서 $j(\neq i, \neq v)$ 로 가는 최단경로 가운데 노드 $v(\neq i)$ 를 거치는 경로의 수이다.

3. 연구방법

3.1 연구대상과 수집

정보교육 관련 석·박사 학위 논문을 분석하기 위해

한국교육학술정보원에서 운영하는 학술연구정보서비스 (<http://www.riss.kr>) 에서 정보교육 관련 학위논문을 수집하였다. 키워드는 ‘정보 교육’, ‘컴퓨터 교육’, ‘ICT 교육’, ‘SW 교육’으로 검색하였으며, 연도는 전체기간으로 설정하였다. 키워드를 더 세부적으로 나누면 수집에서 제외되는 논문이 있을 수 있으므로 정보교육과 유사한 대표적인 키워드를 중심으로 선정하였다. 또한, 컴퓨터 교육은 컴퓨터 소양교육, 컴퓨터과학 교육, 컴퓨터를 교수학습 및 교육 관리에 활용하는 교육을 포함하고 있으므로 정보관련 학과에서 발표되는 논문으로 한정하지 않고 전체학과로 검색하였다. 이클립스를 통해 데이터를 수집하였으며 중복 추출된 논문을 제외한 후 25,773건을 수집하였다.

단계별 논문의 흐름을 파악하기 위하여 1969년부터 1999년까지 제 1기, 2000년부터 2004년까지 제 2기, 2005년부터 2009년까지 제 3기, 2010년부터 2016년까지 제 4기로 구분하였다. 2000년 ‘ICT 교육 운용지침’을 통해 학교에서 정보교육이 활성화된 계기가 되었으므로 2000년 전과 후를 구분할 필요가 있으며, 그 이후 2014년 ‘SW교육 운영지침’이 발표되었으므로 기간의 구분을 5년 단위로 2000년부터 2004년까지, 2005년부터 2009년까지, 2010년부터 2016년까지 설정하였다. 수집된 데이터 총수와 기간별 데이터 수는 <Table 1>과 같다.

<Table 1> Number of analyzed data

Date range		Number of Data
Period 1	1969-1999	2,664
Period 2	2000-2004	7,966
Period 3	2005-2009	7,863
Period 4	2010-2016	7,279
Whole Period		25,772

3.2 분석방법

정보교육 분야 석·박사 학위논문 빅데이터 분석을 위해 R 프로그램을 이용하여 논문명에서 텍스트 분석을 실시하고, 사회연결망 분석을 통해 기간별 중심단어들의 연계성을 알아본 후 정보교육 분야 석·박사 학위논문의 흐름을 분석하였다.

학술연구정보서비스에서 수집된 논문 데이터의 항목은 논문명, 저자명, 대학명, 연도이며. 한글, 한문, 영어

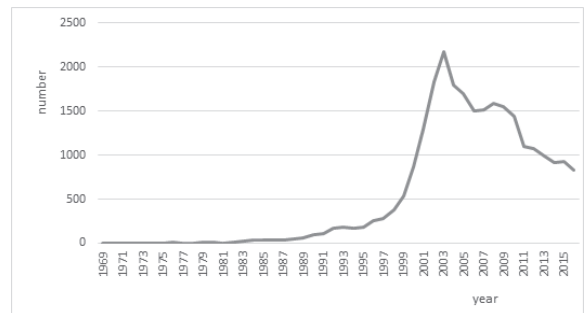
가 혼용되어 있으므로 한문은 한글로 변환한 후 데이터를 정제하였다. 또한, 정보교육과 관련된 복합명사 175건을 선정하여 사전에 추가한 후 처리하였으며, R의 KoNLP 패키지를 사용하여 불용어 및 기호, 숫자 등을 제거하였다. 그리고 논문 제목 마지막에 주로 쓰이는 단어(구현, 개발, 연구, 방안) 등도 조사하여 제거하였다. 이러한 데이터 정제과정을 거친 후 자연어 처리 기술중 형태소 분석 기법을 통해 문장을 최소단위로 쪼개는 작업을 진행하여 단어길이가 두 글자 이상인 명사를 추출하였다.

그리고 추출된 단어들 간의 연결 관계를 파악하기 위해 사회연결망분석을 실시하여 매개중심성의 결과를 통해 중심된 단어를 알 수 있고 중심된 단어들의 연결 관계를 시각화하여 정보교육 분야의 연구방향을 분석하였다.

4. 연구결과 및 분석

4.1 논문현황

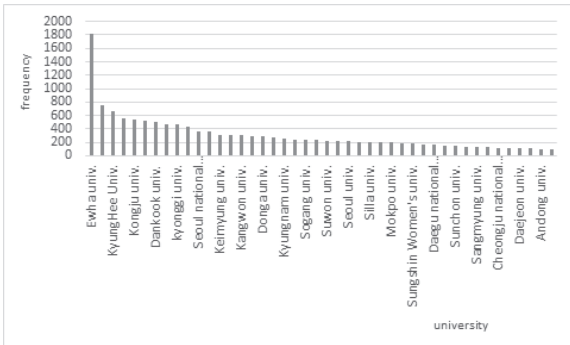
정보교육 관련 석·박사학위논문은 1969년을 시작으로 2000년을 기점으로 논문수가 급증하였다. 2000년 872편에서 2001년 1,298편으로 약 1.5배가 증가한 후, 2003년 2,170편으로 가장 많은 석·박사학위 논문이 발표되었으며 그 이후 발표편수가 조금씩 감소된 후 2013년부터 평균 약 900편정도가 발표되고 있다. 연도별 논문 수는 (Fig. 1)과 같다,



(Fig. 1) Trend in number of data in 1969-2015

정보교육 관련 석·박사학위 논문은 전국 148개 대학에서 산출되었으며 이화여자대학교에서 가장 많고 고려

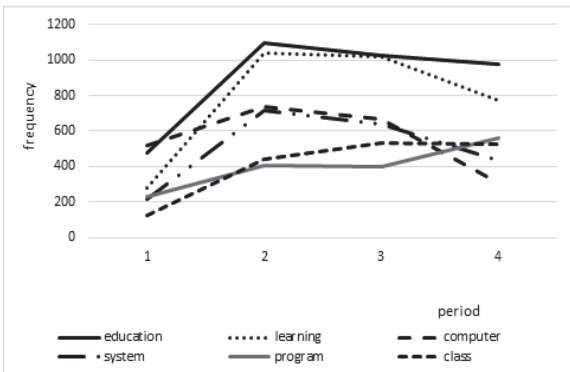
대학교, 경희대학교, 숙명여자대학교, 공주대학교 순이었다. 교육대학교중에서는 서울교육대학교가 발표편수가 가장 많았다. 그중 100편 이상 논문이 발표된 대학은 44개 대학이며, (Fig. 2)와 같다.



(Fig. 2) Number of data trend among the universities

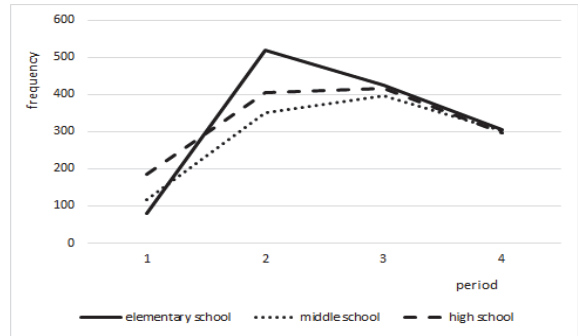
4.2 기간별 단어 분석

수집된 정보교육 분야 석·박사 논문에서 출현빈도가 높은 상위 20개 단어중 모든 기간에서 빈도수가 높은 단어는 기간별로 차이는 있지만 주로 교육, 학습, 컴퓨터, 시스템, 프로그램, 수업, 정보 등이었다. (Fig. 3)의 교육, 학습, 컴퓨터, 시스템, 프로그램, 수업 단어의 기간별 출현빈도를 살펴보면 기간별로 꾸준히 높게 나타나고 있으며, 기간2이후로 빈도수가 증가하여 이후 어느 정도 비슷하게 유지되고 있으므로 기간1부터 기간4까지 관련 연구들이 꾸준히 진행되고 있다는 것을 알 수 있었다.



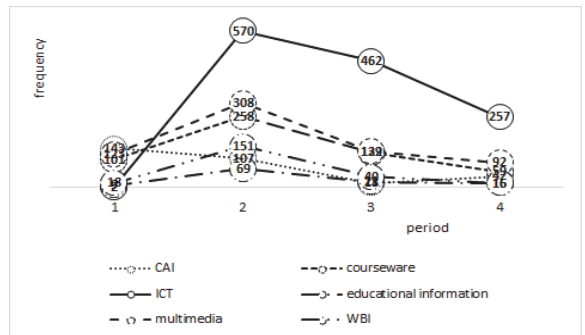
(Fig. 3) Trend in number of data in a given period (education, learning, computer, system, program, class)

또한, (Fig. 4)와 같이 초등학교, 중학교, 고등학교 단어도 지속적으로 높게 나타나는데, 특히 기간2에서 초등학교(518회) 단어는 중학교(350회), 고등학교(404회) 단어보다 높은 출현빈도를 보였으며, 이후 기간3과 기간4에서 서로 비슷한 출현빈도를 나타내었다. 이는 2000년 'ICT 교육 운영지침'의 시행으로 초등학교 1학년부터 주당 1시간씩 컴퓨터 수업을 실시하여 기간2에서 초등학교 관련 논문들이 증가한 것으로 보인다. 또한, 기간1을 제외하고는 기간2부터 초등학교 단어가 중학교, 고등학교 보다 더 많이 등장하였다. 기간별 차이는 있지만 기간1부터 기간4까지 초·중·고등학교 정보교육관련 연구들이 꾸준히 진행되고 있음을 알 수 있었다.



(Fig. 4) Number of data trend in a given period (elementary·middle·high school)

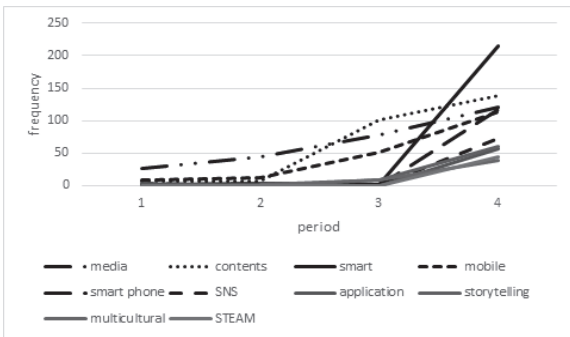
또한, (Fig. 5)와 같이 기간2에서 단어 출현빈도가 높고 기간3, 기간4에서는 점점 단어 출현빈도가 감소되는 현상을 보이는 대표적인 단어들은 CAI, 코스웨어, ICT, 교육정보화, 멀티미디어, WBI 이었다. 이는 2000년대부터 ICT 교육이 강조되고 교육정보화 사업이 한창 진행



(Fig. 5) Words with a high frequency in period 2

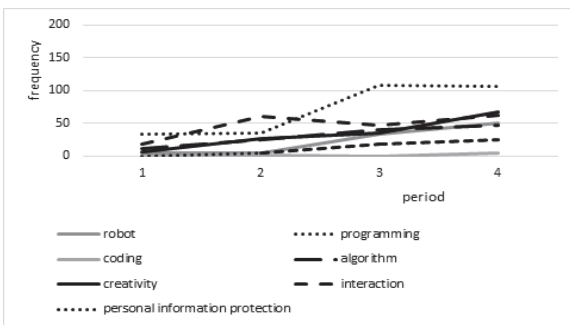
되면서 교육용 소프트웨어 제작에 관심을 갖게 되는 연구가 많아졌다가 2000년 후반부터 점점 모바일 환경과 스마트폰의 등장으로 연구방향이 변경되고 있는 것으로 판단된다.

(Fig. 6)과 같이 기간4에서 단어 출현빈도가 갑자기 상승되는 단어는 스마트 교육 환경에서의 관련 연구 증가로 스마트, 모바일, 스마트폰, SNS, 어플리케이션 단어들과 스토리텔링, 다문화, STEAM 단어들이며, 콘텐츠 단어는 기간3부터 빈도수가 상승되어 기간4에도 활발하게 연구되는 단어이다.



(Fig. 6) Words with a high frequency in period 4

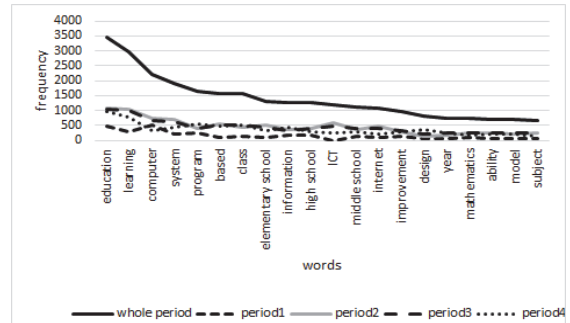
(Fig. 7)과 같이 기간4로 갈수록 빈도수가 점점 증가되지만 단어 출현빈도수는 크지 않고 모든 기간에서 꾸준히 등장하는 단어들은 로봇, 프로그래밍, 코딩, 알고리즘, 창의성, 상호작용, 개인정보보호 단어들이었다.



(Fig. 7) Words with a constant frequency from period 1 to period 4

전체기간에서 단어 출현빈도수가 높은 상위 20개의 단어는 (Fig. 8)과 같고, 기간별 단어빈도수는 <Table

2>와 같다.



(Fig. 8) Top 20 words in terms of a frequency along the period

<Table 2> Frequency of words in each period

period	number	words
Whole Period	over 1,000	education, learning, computer, system, program, base, class, elementary school, information, high school, ICT, middle school, internet
	999~500	improvement, design, grade, mathematics, ability, model, subject, teacher, english, multimedia, awareness, student, didactics, elementary school student, society, curriculum, oneself, case, activity, learner, improvement
	499~300	cyber, elementary, courseware, activation, problem, school, relationship, assessment, environment, unit, method, web-based, game, survey, comparison, youth, factor, satisfaction, infant, digital, information, management, addiction, textbook, characteristics, professor, data, level, CAI, science, kidney, efficiency, attitude, educational, operation, use, academic achievement, building, online, status, lead, Middle school, course, technology, art, information, communication, strategy
	over 500	computer
Period 1	499~200	education, learning, program, system
	199~100	high school, information, CAI, improvement, class, multimedia, middle school, courseware
Period 2	over 500	education, learning, computer, system, ICT, infrastructure, elementary school
	499~200	internet, class, high school, program, information, middle school, multimedia, improvement, courseware, web-based, math, subject, ability, didactics, English, teacher, society
	199~100	model, grade, self, curriculum, instruction, informatization, unit, cyber, elementary, assessment, problem, improvement, WBI, student, method, development, lead, elementary, case, youth, design, level, school, construction, cognition, infant,

		science, environment, resource, activity, learner, virtual, remote, game, management, professor, operational, status, CAI, relationship, comparison, electronic
Period 3	over 500	education, learning, computer, system, class
	499~200	base, ICT, elementary school, high school, program, middle school, internet, improvement, information, subject, model, grade, math, ability, design, elementary school student, math, English, map, teacher
	199~100	curriculum, self, student, awareness, game, addiction, cyber, improvement, school, evaluation, problem, relationship, method, case, activation, activity, unit, social, elementary, learner, environment, youth, web-based, academic achievement, information, communication, satisfaction, comparison, multimedia, operation, courseware, survey, data, textbook, online, led, level, informatization, factor, management, efficiency, art, programming, middle school, digital, current, music, content, professor, development
Period 4	over 500	education, learning, program, class
	499~200	Infrastructure, Information, system, design, computer, elementary, middle school, High school, recognition, improvement, grade, ICT, internet, smart, ability, model, elementary, activity, case
	199~100	factor, student, relationship, digital, Mathematics, English, teacher, elementary, satisfaction, learner, Social, guidance, self, characteristic, curriculum, activation, cyber, enhancement, environment, subject, textbook, comparisons, didactics, service, problem, childhood, content, attitude, art, school, management, target, use, youth, game, running, evaluation, university, media, unit, science, method, smartphone, online, addiction, research, mobile, middle school student, integration, Professor, technology, network, level, strategy, programming, medium, technique, course

4.3 사회연결망 분석

정보교육 분야 석·박사학위 논문에 출현된 단어빈도 수 뿐만 아니라 단어들 간의 연결정도를 분석하여 기간 별로 연구흐름을 좀 더 세부적으로 분석하고자 사회연결망 분석을 실시하였다. 사회연결망 분석은 개인과 집단들 간의 관계를 노드와 링크로서 모델링해 그것의 구조나, 확산 및 진화과정을 계량적으로 분석하는 방법이다.

석·박사학위 논문의 단어들의 네트워크를 파악하기 위해 중심성 여부를 측정된 결과 <Table 3>과 같다. 한 노드에 직접적으로 연결되는 링크정도를 나타내는 연결정도 중심성이 높은 단어는 교육, 학습이며, 그 다음으로 초등학교, 컴퓨터, 기반, 시스템 순이다. 직접적으로

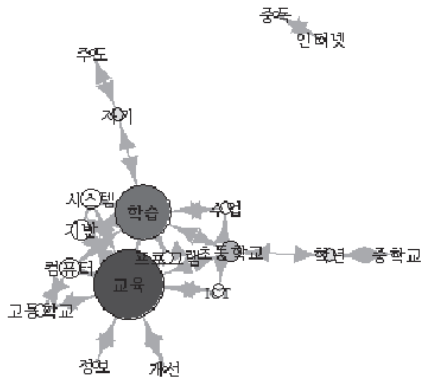
연결된 노드뿐만 아니라 간접적으로 연결된 모든 노드 간의 거리를 합산해 중심성을 측정하는 근접중심성이 높은 단어는 교육, 학습, 초등학교, 컴퓨터, 기반, 시스템, 프로그램, ICT 순으로 연결정도 중심성과 비슷한 결과를 보인다. 한 노드가 연결망 내의 다른 노드들 사이의 최다 경로 위에 위치하는 정도로 중심성을 측정하는 매개중심성이 높은 단어는 교육, 학습, 초등학교, 학년, 컴퓨터, ICT, 수업 순이다. 즉, 교육, 학습, 초등학교, 학년, 컴퓨터, ICT, 수업 단어는 다른 단어들과 직간접적으로 연결된 노드가 많으며 다른 단어들과의 매개역할도 많이 하는 것으로 나타나 이 네트워크에서 중심성이 큰 단어로 볼 수 있다.

<Table 3> Value of degree centrality, closeness centrality and betweenness centrality of each word

words	Degree Centrality	Closeness Centrality	Betweenness Centrality
education	20	0.01724	106.33
learning	16	0.01694	87.33
elementary school	6	0.01562	52
self	4	0.01408	28
grade	4	0.01315	28
computer	6	0.01492	3.67
ICT	4	0.01428	3.67
class	4	0.01408	3
based	6	0.01492	0
system	6	0.01492	0
program	4	0.01470	0
high school	4	0.01408	0
refinement	2	0.01388	0
information	2	0.01388	0
directed	2	0.01176	0
middle school	2	0.01111	0
addiction	2	0.00346	0
internet	2	0.00346	0

전체 데이터의 매개중심성을 도식화하면 (Fig. 9)와 같이 교육 단어를 중심으로 ICT, 개선, 정보, 고등학교, 컴퓨터, 기반, 프로그램, 초등학교, 시스템 단어들이 둘러싸고 있으며, 학습이라는 단어를 중심으로 수업, 초등학교, 프로그램, 기반, 시스템, 자기주도 단어들이 둘러싸고 있다. 이는 이와 같은 단어들과 밀접하게 관련되어 연구가 진행되었음을 알 수 있으므로 컴퓨터교육, 정보교육, ICT교육, 프로그램 교육과 정보교육의 자기주도

학습에 관한 연구도 관심있게 진행된 것으로 생각된다. 초·중학교는 주로 학년별 ICT 수업과 관련된 연구가 진행되었고, 고등학교는 정보교육과 관련된 연구가 이루어진 것으로 판단되며, 학습시스템 관련 연구와 인터넷 중독 연구도 진행된 것으로 보인다.

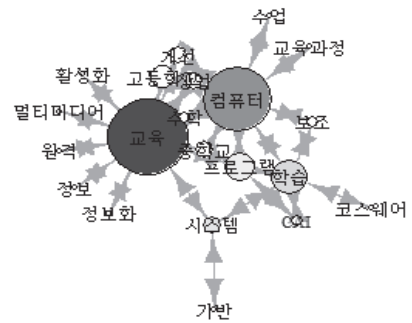


(Fig. 9) Schematic map of closeness centrality in the whole period

기간별로 나누어 매개중심성을 살펴보면 먼저 기간1에서는 교육, 컴퓨터, 학습, 시스템, 프로그램, 고등학교 단어가 매개중심성이 크며, 기간2에서는 학습, 교육, ICT, 초등학교, 컴퓨터, 시스템 단어가 매개중심성이 크게 나왔다. 기간3에서는 학습, 교육, 컴퓨터, 초등학교, ICT, 수업, 시스템, 중학교, 학년 단어가 매개중심성이 크며, 기간4에서는 교육, 학습, 학년 단어가 매개중심성이 크게 나왔다.

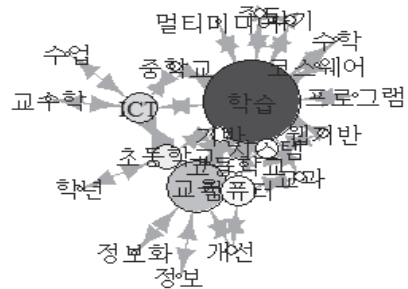
기간1의 매개중심성을 도식화한 (Fig. 10)을 살펴보면, 교육 단어를 중심으로 멀티미디어, 원격, 정보, 정보화, 시스템, 컴퓨터, 고등학교, 활성화 단어들이 둘러싸고 있으며, 컴퓨터 단어를 중심으로 수업, 교육과정, 학습, 프로그램, 개선, 중학교, 고등학교 단어들이 둘러싸고 있다. 이는 이와 같은 단어들과 밀접하게 관련되어 연구가 진행되었음을 알 수 있으며, 코스웨어, CAI, 멀티미디어, 원격 단어의 출현은 'CAI 코스웨어 설계 및 개발', '멀티미디어 코스웨어 개발', '원격교육' 관련 연구가 이 기간에 활발하게 이루어진 것으로 보인다. 특히 수학 교과에서 컴퓨터와 연관된 연구가 활발하게 진행된 것으로 보인다.

기간2의 매개중심성을 도식화한 (Fig. 11)을 살펴보



(Fig. 10) Schematic map of closeness centrality in period 1

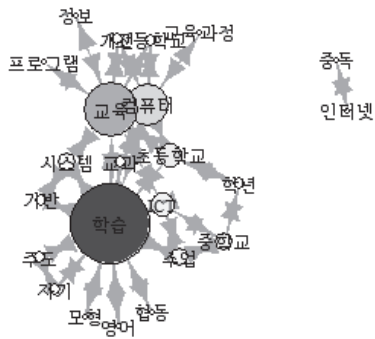
면, 학습 단어를 중심으로 코스웨어, 프로그램, 수업, 자기주도, 멀티미디어, 웹기반, 시스템, 컴퓨터, 교과, 중학교, 고등학교 단어들이 둘러싸고 있으며, 교육 단어를 중심으로 컴퓨터, 고등학교, 개선, 정보화 단어들이, ICT 단어를 중심으로 초등학교, 중학교, 교과, 수업 단어들이 둘러싸고 있다. 이는 이와 같은 단어들과 밀접하게 관련되어 연구가 진행되었음을 알 수 있으며, 기간1에 비해 ICT, 초등학교 단어의 매개중심성이 높아졌고, 웹기반, 초등학교, 교과 단어들이 새롭게 출현하였다. 이 기간에 ICT 교육, 웹기반 학습과 관련된 '초·중학교 ICT 수업 및 교과활용수업', '웹기반 학습시스템', '교육정보화' 등의 연구가 이루어진 것으로 보인다.



(Fig. 11) Schematic map of closeness centrality in period 2

(Fig. 12)의 기간3에서는 학습 단어를 중심으로 협동, 자기주도, 모형, 기반, 시스템, 교과, ICT, 수업, 영어 단어들이 둘러싸고 있으며, 교육 단어를 중심으로 교과, 시스템, 정보, 프로그램, 고등학교, 컴퓨터, 개선 단어들이, 컴퓨터 단어를 중심으로 초등학교, 교과, 교육, 개선,

교육과정 단어가 둘러싸고 있다. 이는 이와 같은 단어들과 밀접하게 관련되어 연구가 진행되었음을 알 수 있으며, 협동, 영어, 모형 단어가 새롭게 출현하였다. 정보교육에서 다양한 수업모형을 적용한 연구들과 협동학습, 자기주도학습 관련 연구와 인터넷 중독 관련 연구가 이루어진 것으로 보이며, 특히 영어 교과에서 컴퓨터와 연관된 연구가 활발하게 진행된 것으로 보인다.

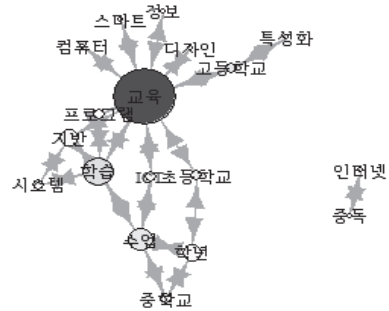


(Fig. 12) Schematic map of closeness centrality in period 3

(Fig. 13)의 기간4에서는 교육 단어를 중심으로 고등학교, 디자인, 정보, 컴퓨터, 스마트, 컴퓨터, 프로그램 단어가 둘러싸고 있으며, 학습 단어를 중심으로 수업, 기반, 시스템, 프로그램 단어가, 학년 단어를 중심으로 초등학교, 중학교 단어가 둘러싸고 있다. 이는 이와 같은 단어들과 서로 밀접하게 관련되어 연구가 진행되었음을 알 수 있으며, 스마트, 디자인, 특성화 단어가 새롭게 출현하였다. 스마트 기기 환경에서의 교육 방안과 디자인 교육 및 개발 관련 연구, 특성화 고등학교 대상 관련 연구가 이루어진 것으로 보인다.

정리하면 정보교육분야 석·박사학위 논문의 방향은 전반적으로 컴퓨터교육, 정보교육, ICT교육, 프로그램 교육과 정보교육의 자기주도학습에 관한 연구가 관심있게 진행된 것으로 보이며, 초·중학교는 주로 학년별 ICT 수업, 고등학교는 정보교육 관련 연구에 초점을 이룬 것으로 판단된다. 또한, 학습시스템 및 인터넷 중독 연구도 이루어졌다. 기간별 연구흐름은 기간1에서는 컴퓨터보조학습, 멀티미디어 개발, 원격교육, 기간2에서는 ICT 교육, 웹기반 학습, 교육정보화, 기간3에서는 정보교육관련 수업모형, 기간4에서는 스마트교육 등의 연구

흐름의 특징을 보인다고 할 수 있다.



(Fig. 13) Schematic map of closeness centrality in period 4

5. 결론 및 제언

사회 연결망 분석은 개체간의 관계를 기술하고 구조화하여 이들 간의 관계를 분석하는데 사용되며, 데이터만으로는 그 모습을 파악하기 어려운 연결 구조를 한눈에 파악할 수 있도록 도와준다[3,7]. 이런 측면에서 사회 연결망분석을 이용한 연구가 빅데이터 분석에 사용되고 있다.

이에 본 논문에서는 정보교육 관련 석·박사 학위논문을 수집하여 텍스트분석과 사회연결망분석을 실시하여 논문의 흐름을 분석하였다.

정보교육 관련 석·박사학위논문은 1969년을 시작하여 2000년을 기점으로 논문수가 급증하였으며, 2003년에 가장 많은 논문이 발표된 후 조금씩 감소되다가 2013년부터 평균 편수가 유지되고 있다.

분석결과 첫째, 모든 기간에서 출현빈도수가 높게 나오거나 지속적으로 나오는 단어가 있다. 모든 기간에서 출현빈도가 높은 단어는 교육, 학습, 컴퓨터, 시스템, 프로그램, 수업, 정보 등이며, 이러한 단어들과 관련된 연구들이 지속적으로 진행되고 있다고 판단할 수 있다. 초등학교, 중학교, 고등학교 단어도 지속적으로 높게 나타나므로 초·중·고등학교 정보교육관련 연구들이 꾸준히 진행되고 있음을 알 수 있다. 또한, 단어 출현빈도수는 크지 않고 모든 기간에서 꾸준히 등장하는 단어들은 로봇, 프로그래밍, 코딩, 알고리즘, 창의성, 상호작용, 개인정보

호 단어들이 많이 나올 수 있어 이러한 단어들과 관련된 연구들이 지속적으로 진행되고 있다고 판단할 수 있다.

둘째, 각 기간마다 출현빈도가 높아진 단어들이 있다. 기간2에서는 CAI, 코스웨어, ICT, 교육정보화, 멀티미디어, WBI이며, 기간4에서 스마트, 모바일, 스마트폰, SNS, 어플리케이션 단어들과 스토리텔링, 다문화, STEAM 이며, 기간3에서는 콘텐츠 단어의 출현 빈도수가 상승되어 기간4에도 활발하게 연구되어졌다. 이는 2000년대부터 ICT교육이 강조되고 교육정보화사업이 한창 진행되면서 교육용 소프트웨어 제작에 관심을 갖게 되는 연구가 많아졌다가 2000년후반부터 점점 모바일 환경과 스마트폰의 등장으로 연구방향이 변경되고 있는 것으로 판단된다.

셋째, 사회연결망 분석 결과 기간별 연구흐름의 특징이 있다. 기간1에서는 컴퓨터보조학습, 멀티미디어 개발, 원격교육, 기간2에서는 ICT 교육, 웹기반 학습, 교육정보화, 기간3에서는 정보교육관련 수업모형, 기간4에서는 스마트교육 등의 연구흐름의 특징을 보인다고 할 수 있다.

그러므로 IT 기술발전과 초·중·고등학교 정보교육과정 변화에 민감하게 정보교육 석·박사학위 논문주제가 변화되었음을 알 수 있었다. 앞으로 기간4에서 출현빈도가 높아진 스마트, 모바일, 스마트폰, SNS, 어플리케이션, 스토리텔링, 다문화, STEAM과 관련된 연구가 지속될 것으로 예측하며, 로봇, 프로그래밍, 코딩, 알고리즘, 창의성, 상호작용, 개인정보보호와 관련된 주제도 꾸준히 연구될 것으로 예측된다.

본 논문에서는 논문제목에 중점으로 연구흐름을 분석하였으나 요약정보를 포함하여 사회연결망 분석을 실시해보면 제목에서 나타나지 않은 세부 연구내용들이 추출될 수 있어 보다 세분화된 연구방향을 분석할 수 있을 것이다. 그리고 연구자 정보를 포함하여 사회연결망 분석을 실시해보면 세부연구 주제별 연구자들의 사회연결망 분석이 가능할 것이다. 또한 국외 석·박사학위 논문을 포함하여 분석한다면 국내외 정보교육 연구방향을 비교 분석할 수 있을 것이다.

참고문헌

[1] Bae, KyuYong, Park, JuHyun, Kim, JeongSeon, &

- Lee YungSeop(2013), Analysis of the abstracts of research articles in food related to climate change using a text-mining algorithm, *Journal of the Korean Data & Information Science Sociert*, 24(6), 1429-1437.
- [2] Bae, SungHoon, Kim, JoonHyun, Kang, SangGyu, Yoon, JinSun, Shin, GwangMin, Kim, MinGwan, Lee, JeongWoo, & Jan, ChangHee(2016), Social Network Analysis for Carbon Nano - Materials Technology Research Trends, *Proceedings of the Korean Institute of Industrial Engineers summer conference, 2016*(4), 5440-5449.
- [3] Choi, KyoungHo, & Yoo, JinAh(2015), A reviews on the social network analysis using R, *Journal of the Korea Convergence Society* 6(1), 77-83.
- [4] Hampton, K., & Wellman, B. (2001). Long distance community in the network society: Contact and support beyond Netville. *American Behavioral Scientist*, 45 476-495.
- [5] Hong, Jisuk, & Oh, Ickkeun(2016), Image difference of before and after an incident using social big data analysis : Focusing on a ramp return of "K" airline, *International Journal of Tourism and Hospitality Research*, 30(6), 119-133.
- [6] Hwang, DeukYoung, & Kim. JinMook(2015), A study of Big-data analysis for relationship between students, *Convergence security journal*, 15(4), 113-119.
- [7] Hwang, DongRyul, Moon, SinJung, & Hwang, GoEun(2016), Social Network Analysis of Dance Studies utilizing Big Data, *The Korean Journal of Physical Education*, 55(1), 611-621.
- [8] Ihaka, Ross, & Robert Gentleman(1996), R: a language for data analysis and graphics, *statistics* 5.3, 299-314.
- [9] Jang, YoonJae, & Cho, SungGyeong(2014), A Comparative Analysis of Data Gathering and Sampling Methods for Social Data, *Journal of Institute for Social Sciences*, 25(2), 3-25.
- [10] Jeong, YoungSik, & Yoo, JeongSoo(2015),

Software Education, Cmass.

- [11] Kang, Ae Tti, & Kang, YoungOk(2015), Location Inference of Twitter Users using Timeline Data, *The Journal of Korea Spatial Information Society*, 23(2), 69-81.
- [12] Kim, HyoJeong(2006), Social Network Analysis of Spontaneous Teacher Network, *The Journal of Educational Administration*, 24(2), 173-200.
- [13] Kim, SeongHee, & Chang, RhoSa(2010), The Study on the Research Trend of Social Network Analysis and the its Applicability to Information Science, *Journal of the Korean Society for Information Management*, 27(4), 71-87.
- [14] Kim, Yong Hak, & Yu, So Young(2013), A Comparative Study of Comparative Studies of Korea and Japan: Co-word Analysis in Social Sciences and Humanities, *social science review*, 44(1), 25-45.
- [15] Kim, YongHak(2014), Social Network Anlaysia, Park Young Sa.
- [16] Lee, JuYeon, & Park, YooHyun(2016), Scocial Netword Analysis of author's Interest area in *Journals about Computer, JKIIICE*, 20(1), 193-199.
- [17] Lee, TaeWook, & Choi, HyunJong(2016), Education of Information Hanbit academy.
- [18] Park, Sunju(2015), Topic Analysis of SW education textdata using R, *Journal of the KAIE*, 19(4), 517-524.
- [19] Park, Sunju(2016), The Analysis on the KAIE Articles using Social Network Analysis, *Journal of the KAIE*, 20(6), 543-551.
- [20] Seo, JungA(2016), Analyzing the Destination Image of Daegu from Online Content through Social Network Anlaysia, Keimyung University PhD. thesis.
- [21] Utz, S. (2009). The (potential) benefits of campaigning via social network sites. *Journal of Computer Mediated Communication*, 14(2), 221-243.

저자소개

박 선 주



1995 전남대학교 전산통계학과 (이학박사)
 2003 George Mason University 객원교수
 1996~현재 광주교육대학교 컴퓨터교육과 교수
 관심분야 : 컴퓨터교육, SW교육, 앱개발, 빅데이터
 E-mail : sjpark@gnue.ac.kr