

# 키워드 네트워크 분석을 활용한 세계 크루즈산업 연구동향

장세은\* · † 이수호

\* 한국해양대학교 영어영문학과 교수, † 한국해양대학교 국제무역경제학부 교수

## Research Trends in Global Cruise Industry Using Keyword Network Analysis

Se-Eun Jhang\* · † Su-Ho Lee

\* Department of English Language and Literature, Korea Maritime and Ocean University, Busan 606-791, Korea

† Division of International Trade and Economics, Korea Maritime and Ocean University, Busan 606-791, Korea

**요 약** : 세계적으로 해양산업은 크루즈산업에 많은 관심과 연구가 대폭적으로 이루어지고 있고 우리나라도 미래의 잠재력 있는 국가동력 산업의 하나로 인식하고는 있으나 크루즈산업의 연구동향 분석 연구는 국내외에 활발히 이루어지고 있지 않다. 따라서 우선 세계 크루즈산업을 연구하고 이해하기 위해서는 최근에 다양한 산업에 대한 연구동향을 분석한 방법을 활용하여 크루즈산업에도 적용할 필요가 있다. 본 논문의 목적은 외국 유명저널에 발표된 크루즈산업과 관련된 학술논문에서 제시하고 있는 키워드와 논문을 매개로 한 키워드 네트워크를 구축하여 복잡계의 네트워크 분석에서 사용하는 연결 중심성과 매개 중심성 분석방법으로 시대별로 나누어 시각화하여 살펴봄으로써 세계 크루즈산업의 연구동향을 심층적으로 관찰하여 논의하는 것이다. 본 연구에서 제시된 키워드 빈도는 Zipf의 법칙을 따르고 노드의 연결정도는 멱함수 분포를 보여주고 있어 언어네트워크에서 분석하는 키워드 네트워크와 동일함을 관찰한다. 연구방법론으로는 키워드 네트워크 분석을 위하여 사회연결망 프로그램인 넷마이너 4.0을 사용하여 여러 가지 중심성 측정방법 중 키워드 상위 20개의 빈도순위를 비교하여 빈도순위와 가장 가까운 중심성 측정방법을 선택하여 크루즈산업의 연구동향을 분석한다. 특히 크루즈산업의 연구동향이 연도 기간별로 어떠한 변화를 가져왔는지를 살펴보기 위해 2000년 이전과 2000년 이후로 크게 대별하여 나누고 2000년 이후에는 5년 주기로 각 기간별 공통적으로 나타나는 연결 중심성이 높은 최상위의 키워드인 *cruise*와 *tourism* 노드를 중심으로 매개 중심성이 높은 것들의 키워드 네트워크를 시각화하여 논의한다. 연구결과에서 흥미롭게도 2010-2014의 기간에 새로운 노드로 *China*가 등장하여 최상위의 키워드들을 연결하고 있는 것은 최근 급성장하고 있는 중국의 크루즈산업의 발전 양상을 보여준다. 그러므로 본 연구에서 사용하는 키워드 네트워크 분석은 각 연도 기간별 네트워크의 다른 종류의 숫자와 크고 작은 중심축 군집 네트워크의 숫자의 증감뿐 아니라 중심축 군집 네트워크의 중심에 있는 키워드 간의 연결 분석을 용이하게 해주어 기간별 연구동향을 파악하는데 유용한 방법임을 확인할 수 있었다.

**핵심용어** : 키워드, 네트워크, 연결 중심성, 매개 중심성, 넷마이너, 크루즈산업, 연구동향

**Abstract** : This article aims to explore and discuss research trends in global cruise industry using keyword network analysis. We visualize keyword networks in each of four groups of 1982-1999, 2000-2004, 2005-2009, 2010-2014 based on the top 20 keyword nodes' degree centrality and betweenness centrality which are selected among four centrality measurements, comparing them with frequency order. The article shows that keyword frequency collected from 240 articles published in international journals is subject to Zipf's law and nodes degree distribution also exhibits power law. We try to find out research trends in global cruise industry to change some important keywords diachronically, visualizing several networks focusing on the top two keywords, *cruise* and *tourism*, belonging to all the four year groups, with high degree and betweenness centrality values. Interestingly enough, a new node, *China*, connecting the top most keywords, appears in the most recent period of 2010-2014 when *China* has emerged as one of the rapid development countries in global cruise industry. Therefore keyword network analysis used in this article will be useful to understand research trends in global cruise industry because of increase and decrease of numbers of network types in different year groups and the visual connection between important nodes in giant components.

**Key words** : keywords, network, degree centrality, betweenness centrality, NetMiner, cruise industry, research trends

### 1. 서론

세계의 해양강국들은 미래의 해양산업은 크루즈관광이 주도한다고 믿고 크루즈산업에 많은 관심과 연구가 대폭적으로 이루어지고 있고 우리나라도 미래의 잠재력 있는 국가동력산업의 하나로 인식하고는 있으나 크루즈산업의 연구동향 분석 연

구는 국내외에 활발히 이루어지고 있지 않다(Cho et al., 2012). 우선 세계 크루즈산업을 연구하고 이해하기 위해서는 최근에 다양한 산업에 대한 연구동향을 분석한 방법을 활용하여 크루즈산업에도 적용할 필요가 있다. 지금까지 외국 유명저널에 발표된 크루즈산업과 관련된 학술논문에서 제시하고 있는 키워드와 논문을 매개로 한 키워드 네트워크를 구축하여

\* 연희원, jhang@kmou.ac.kr 051)410-4595

† Corresponding author : 연희원, leesh@kmou.ac.kr 051)410-4404

복잡계의 네트워크 분석에서 사용하는 중심성 분석방법으로 시대별로 나누어 시각화하여 살펴봄으로써 세계 크루즈산업의 연구동향을 심층적으로 관찰할 수 있을 것이다.

본 연구에서는 기존의 선행연구에서 사용하고 있는 키워드 네트워크 분석을 위하여 사회연결망 프로그램인 넷마이너(Cyram, 2013) 4.0을 사용하여 여러 가지 중심성 측정방법 중 키워드 상위 20개의 빈도순위를 비교하여 빈도순위와 가장 가까운 중심성 측정방법을 선택하여 크루즈산업의 연구동향을 분석한다. 특히 크루즈산업의 연구동향이 연도 기간별로 어떠한 변화를 가져왔는지를 살펴보기 위해 2000년 이전과 2000년 이후로 크게 대별하여 나누고 2000년 이후에는 5년 주기로 각 기간별 공통적으로 나타나는 연결 중심성(degree centrality)이 높은 최상위의 키워드인 *cruise*와 *tourism* 노드를 중심으로 매개 중심성(betweenness centrality)이 높은 것들의 키워드 네트워크를 시각화하여 논의한다.

본 연구에서 사용하는 키워드 네트워크 분석은 각 연도 기간별 네트워크의 다른 종류의 숫자와 크고 작은 중심축 군집 네트워크의 숫자의 증감뿐 아니라 중심축 군집 네트워크의 중심에 있는 키워드 간의 연결 분석을 용이하게 해주어 기간별 연구동향을 파악하는데 유용한 방법임을 보여주고자 한다.

## 2. 선행연구

본 논문과 관련 있는 국내외 연구논문의 키워드 네트워크를 활용한 학술분야의 연구 경향 또는 산업분야의 동향을 연구한 최근의 선행연구를 살펴본다. 우선 학술분야의 연구 경향에 대한 선행연구부터 살펴보자. Lee and Kang(2011)은 2004년부터 2010년까지 사회학 분야의 학술논문 및 저자정보에 대한 분석을 수행하였으며 사회학의 동향을 분석하기 위하여 사용한 네트워크 도구는 연결 중심성과 매개 중심성이었다. Jang et al.(2012)는 보건의료분야 연구주제 중 암유전자 분야의 지식지도를 구축하고 국내외 연구주제 네트워크를 시계열적으로 분석하였다. 논문의 키워드를 대상으로 사회네트워크 분석인 가중치가 부여된 연결 중심성을 적용하여 연구주제 네트워크를 구축하였으며, 이를 1981년부터 2010년까지 5년 구간별로 나누어 연구하였다. Ryu and Hyun(2013)은 행정학에서 복지 분야의 최근까지의 연구경향을 살펴보았는데 언어네트워크 방법론을 사용한 연구키워드의 중심성 분석을 사용하여 복지 분야 연구경향은 2000년 이후 복지 관련 연구와 이론을 적용하는 연구가 증가함을 보여주었다. Yang(2013)은 민간경비 분야에 대한 연구 경향을 살펴보고자 대상의 경향을 분석하는 연구방법인 언어 네트워크 분석 방법 중 연결 중심성과 근접 중심성(closeness centrality)을 활용하여 분석하였다. Kwon(2014)은 학술논문의 키워드 출현순서에 따른 저자키워드의 특징을 분석한 후 네트워크 분석 방법의 연결 중심성, 근접 중심성, 매개 중심성, 위세 중심성, 그리고 네트워크의 구조적 공백성의 효과 크기와 같은 지수를 사용하여 저자 키워드의 중요도를 측정하였다. Kwon et al.(2014)은 네트워크 분석 기

법을 활용하여 한의학 분야 연구동향 분석을 연결 중심성과 매개 중심성을 사용하여 연도별 비교를 통해서 네트워크의 모습과 연구 중심이 어떻게 변화했는지 연구하였다.

다음으로 산업의 동향을 분석한 연구를 살펴보자. Kho et al.(2013)은 최근 기술경영분야의 주요 해외 저널에 게재된 논문의 키워드를 가지고 네트워크 컴포넌트 분석 및 중심성 분석을 중심으로 키워드 네트워크 구조를 분석하였다. 특히 중심성 분석을 통하여 혁신, 연구개발, 특허, 예측, 기술이전, 기술, 중소기업 등의 키워드가 연결 중심성, 매개 중심성, 근접 중심성이 높음을 확인하였다. Jeong and Kim(2012)은 2010년 2월에 발표된 과학기술 미래비전의 내용 중 건설 분야의 미래 주요 키워드 들을 사용하여 네트워크의 연결 중심성, 매개 중심성, 근접 중심성 분석을 하였다.

외국의 연구 사례를 살펴보면 Borgatti et al.(2009)에서는 지난 수십 년 간 네트워크 연구의 폭발적인 증가가 물리학과 사회학분야에서 일어났으며 특히 사회학자들에게 사회현상을 설명할 수 있는 중요한 도구였음을 보여주었다. 따라서 이 연구는 사회학 분야의 연구를 전반적으로 살펴보고 네트워크 현상과 관련하여 논의하였다. Park et al.(2012)는 학문의 키워드는 한 분야에만 국한된 것이 아니라 학제 간 융합으로 나아가는 추세라고 판단하고 컴퓨터과학 분야의 키워드들이 어떤 방식으로 확산되는지를 1970년부터 2010년까지 5년 단위로 키워드의 변화양상을 연구하였다. Chaurasia and Tiwari(2014)는 미국 뉴욕에서 일어난 911테러와 관련하여 테러리스트들의 그룹과 이들의 관계를 찾기 위하여 네트워크 분석도구인 연결 중심성, 매개 중심성, 근접 중심성 및 고유벡터(Eigenvector)를 사용하였다. 그러나 기존 연구의 한계점으로는 선행연구에 크루즈산업 연구논문을 대상으로 키워드 네트워크를 분석한 논문이 없다는 점이지만 이점이 오히려 본 논문이 가지는 의의가 된다고 할 수 있다.

## 3. 자료 및 연구방법

### 3.1 자료 및 네트워크의 일반 특성

분석 대상 학술지는 영어로 작성된 국제 유명저널을 대상으로 세계 최대의 학술출판사인 Elsevier에서 발행하는 학술저널원문제공 사이트인 Science Direct와 Impact Factor 값을 가지고 있는 저널의 원문을 제공하는 Wiley사의 Wiley-Blackwell와 Springer에서 제공하는 Springer e-journal을 사용하여 'cruise industry'라는 검색어를 넣어 출력되는 논문을 검색하였다. 추출된 논문 가운데 논문제목과 키워드를 살펴보고 크루즈산업 관련 논문만 수집한 결과 1982년부터 2014년까지 출판된 학술논문이 총 240개가 선정되었다.

본 연구의 목적을 위해 저자들이 준 키워드를 대상으로 세계 크루즈산업 동향의 네트워크 분석을 위해 논문의 출판연도, 논문제목, 논문저자, 논문의 키워드 순으로 목록을 정하여 엑셀에 저장하였다. 또한 네트워크 분석을 위한 키워드 교정

작업으로 단수와 복수에 대한 키워드 클렌징(keyword cleansing) 처리를 하였다. 예를 들면, *cruise ship/cruise ships, cruise/cruises, cruise line/cruise lines*와 같이 단수형과 복수형이 있을 경우 단수형으로 통일하고 키워드의 유사성 행렬을 작성하였다. 키워드 유사성 행렬은 키워드×키워드 간 일원(1-mode)행렬로 변환하여 분석하였다. “유사도 계수는 키워드 간의 통계적인 연관성을 잘 평가할 수 있는 코사인(Cosine) 계수를 주로 사용한다. 어떤 개체(case)들을 유사한 집단으로 분류하고자 하는 경우, 개체들이 얼마나 유사한지 또는 유사하지 않는지를 측정하는 개념 개체가 지니고 있는 특성변수들의 값을 이용하여 유사성을 측정하며, 측정 결과는 유사도(similarity index)로 나타난다(이수상, 2013).”

유사도는 -1부터 1까지의 값을 보여주며, -1은 서로 반대되는 경우이고, 0은 벡터들이 독립일 때, 1은 서로 같은 경우를 의미한다. 코사인계수의 계산 공식은 아래 (1)과 같다.

$$= \frac{\sum_{k=1}^n C_{ik}C_{jk}}{\sqrt{\sum_{k=1}^n C_{ik}^2} \sqrt{\sum_{k=1}^n C_{jk}^2}} \quad (1)$$

위의 (1)의 코사인계수를 사용하여 논문의 고유 번호와 각 저자의 키워드로 만들어진 이원모드(2-mode)자료는 키워드만으로 구성된 일원행렬로 바꾼다. 변환 방법은 넷마이너에서 Transform >> Mode >> 2-mode Network 옵션을 사용하여 변환한 후 각 단어에서 할당된 1을 제거하기 위하여 Transform >> Value >> Diagonal을 실행하여 대각선을 0으로 변환하여 올바른 일원모드를 만든다.

이와 같은 방법을 각 시기별 데이터에 적용하여 넣은 결과를 가지고 정리한 것이 아래 Table 1이다. 여기서 논문 수, 키워드 표제어(lemma) 수, 키워드 빈도수를 연도 기간별로 조사하였다.

Table 1 Distribution of Numbers of Articles and Keywords

Number\Years	All	1982-1999	2000-2004	2005-2009	2010-2014
Articles	240	31	46	100	63
Keyword Lemma	724	100	143	298	261
Keyword Frequency	1,111	132	212	442	325

크루즈산업의 연구동향이 연도 기간별로 어떠한 변화를 가져왔는지를 살펴보기 위해 2000년 이전과 2000년 이후로 크게 대별하여 나누고 2000년 이후에는 5년 주기로 세 개 기간으로 나누어 총 네 개 기간으로 나누었다. 2000년 이후 5년 주기로 자료를 구분한 이유는 Jang et al.(2012)와 Park et al.(2012)에서와 같이 일반적으로 보건의료와 산업의 변화가 5년을 기점으로 크게 변화되는 경향이 있고 또한 일반영어의 시대적 언어전이현상이나 변화 발달과정을 살펴볼 때 코퍼스기반 연구(Corpus of Contemporary American English, 1990-1994, 1995-1999, 2000-2004 등, <http://corpus.byu.edu/coca/>)에서도

시기별로 변화의 양상을 파악하는데 도움을 주기 위하여 5년 주기로 구분하였다. 위의 Table 1에서 보는 바와 같이 논문 수는 점차적으로 증가하다가 마지막 기간에서 줄어든다. 마찬가지로 키워드 표제어 수와 빈도수도 꾸준히 증가하다가 마지막 기간에서 감소한다. 한편 전체 자료를 모든 네트워크에서는 논문 수와 키워드 빈도수는 각 연도 기간에서 합한 수와 동일하지만 키워드 표제어 수는 그렇지 않다. 그 이유는 연도별로 같은 키워드가 반복적으로 사용된 것은 빈도수의 합계로 나타나지만 표제어 수는 사용된 키워드의 종류만 언급하므로 점차 줄어들기 때문이다.

크루즈 산업 연구 분야의 논문 키워드를 Fig. 1에서 볼 수 있듯이 빈도수로 나열해 보면 Zipf(1949)의 법칙을 따르는 것을 알 수 있다. Zipf의 법칙에 따르면 자연언어에서 단어의 사용빈도는 순위에 반비례한다. 예를 들어, 키워드 빈도에서 가장 많이 등장한 *cruise, tourism, cruise tourism, cruise ship, cruise industry, satisfaction, Caribbean*과 같은 상위 7위내에 나타나는 키워드는 다른 키워드와 비교해서 빈도수에서 큰 차이를 보인다.

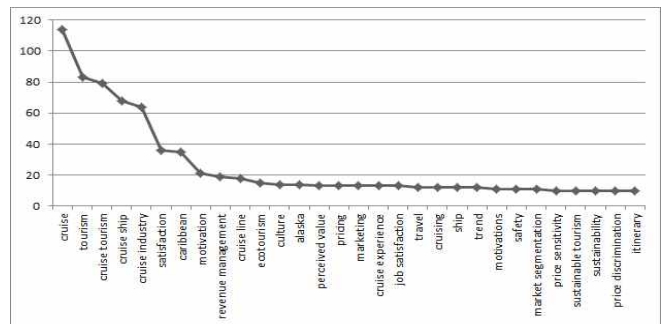


Fig. 1 Zipf's Law Distribution of Keyword Network Frequency in Cruise Industry

그러면 과연 이러한 키워드의 빈도가 키워드 네트워크상에서는 어떤 분포를 나타낼까? Zipf의 법칙이 일종의 멱함수(power law)이므로(Newman, 2005) 네트워크 연결정도가 멱함수 분포를 따르는지 살펴보기 위하여 넷마이너에서 연결정도를 측정하는 log-log 그래프를 그려본 결과 흥미롭게도 아래 Fig. 2가 보여주는 바와 같이 멱함수의 결과를 얻었다.

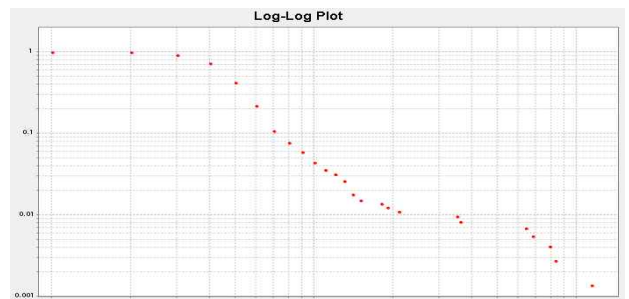


Fig. 2 Power Law Degree Distribution of Keyword Network in Cruise Industry

이러한 멱함수 분포가 의미하는 바는 특정한 몇 개의 키워드가

다른 키워드에 비해 지속적으로 많은 연구자들에게 선정되고 연결되어서 이 주요 키워드들의 조합으로 새로운 연구 분야가 확산된다는 것을 말한다. 또한 Liang et al.(2009)의 연구 결과와 같이 코퍼스 어휘의 빈도 분포는 위의 Fig. 2와 같이 멱함수와 유사하게 나타나기 때문에 언어연구에서 사용하는 키워드 네트워크 분석방법을 여기에 응용할 수 있음을 보여준다.

다음으로 연도 기간별로 키워드 네트워크의 특성을 파악하기 위한 연결(Link), 밀도(Density), 평균연결도(Average Degree), 큰/작은 중심축 집단(Giant/Small Component), 선형(Sting), 과당(Clique)에 관한 일반 특성을 살펴보자.

Table 2 General Characteristics of All and Years Networks

	All	1982-1999	2000-2004	2005-2009	2010-2014
Links	1,894	218	326	693	679
Density	0.007	0.044	0.032	0.016	0.02
Average Degree	2.62	2.18	2.28	2.31	2.60
Giant/Small Components	1/0	1/0	2/0	2/2	1/0
Sting	25	7	10	17	16
Clique	10	4	7	5	1

Table 2에서 연결은 모든 노드의 쌍이 이루고 있는 연결 통로를 말하는데 각 연도별로 지속적인 증가추세를 보인다. 밀도는 전체 노드 간 생성 가능한 링크 수 대비 실제 링크 수를 측정하는 것이다. 1982-1999년에 가장 높고 계속 감소하다 2010-2014년에 다시 증가하였다. 평균연결도는 한 노드가 연결되어 있는 이웃 노드 수의 전체 평균을 의미한다. 연도 기간별로 지속적인 증가세를 보이며 노드들의 연결이 많아진다. 큰/작은 중심축 집단은 연결고리가 끊이지 않는 노드들의 집합을 말하는데 큰 중심축 네트워크는 2000년이 넘어서는 한 개에서 두 개로 증가되다가 2005-2009 기간에는 작은 중심축 네트워크가 두 개가 새로 만들어지다가 마지막 연도 기간에서만 하나로 흡수 통합되었다. 이것은 큰 중심축 하나에 연결되어 있던 노드들이 이전에 없었던 노드 집단들과 연결되면서 두 개의 독립적인 큰 중심축으로 나누어지다가 시간에 흐름에 따라 크루즈 산업의 주요 노드가 계속적으로 생겨나 여러 개의 작은 중심축을 구성하다가 최근에 와서는 하나로 통합되는 패턴을 보여준다. 중심축 노드 집단 이외에 노드의 유형은 두 가지로 더 나누어진다. 선형은 일렬로 늘어선 형태의 노드를 의미하며 노드 상호 간의 연결은 약하다고 볼 수 있다. 과당은 모든 노드들이 직접 연결된 하위 네트워크를 말한다. 적은 수의 노드들이지만 상호 간에 연결이 강한 지역적인 집단들로서 신생하는 키워드의 그룹으로 볼 수 있다. 선형과 과당도 연도 기간별로 점차 늘어나다가 줄어드는 추세를 보이며 특히 과당은 마지막 기간에서는 한 개만 나타나서 과거에 새롭게 늘어난 것들이 큰 중심축 집단에 많이 포함된 것으로 보인다.

3.2 연구방법

본 연구에서는 키워드 네트워크 분석을 활용하여 사회연결

망 프로그램인 넷마이너를 사용하였다. 또한 세계 크루즈산업의 동향을 파악하는데 어떤 중심성 측정방법이 가장 적절한지 알아보기 위하여 다섯 가지 중심성 측정방법 (연결 중심성, 매개 중심성, 근접 중심성, 고유벡터)을 사용하여 서로 비교하였다. 우선 키워드 빈도 상위 20개 키워드를 선택하여 각 중심성이 차지하는 순위의 변동을 빈도와 비교하였다.

상위 20위 키워드 빈도와 다섯 가지 중심성의 순위 비교는 아래 Fig. 3과 같다.

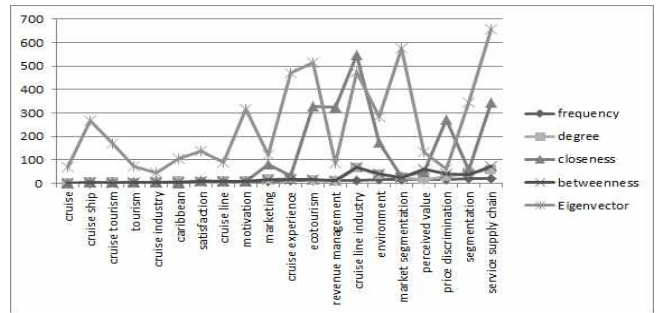


Fig. 3 Comparison of Ranking of the Top 20th Keyword Frequency with That of Five Different Centrality

위의 Fig. 3의 그래프에서 보는 바와 같이 키워드의 빈도가 보여주는 순위 20위와 크게 벗어나는 순위를 가진 중심성 측정방법은 근접 중심성(100위 이상이 6개)와 고유벡터(100위 이상이 6개)이다.

한편, 연결 중심성과 매개 중심성은 빈도 순위와 유사하게 보이므로 더 자세하게 알아보기 위하여 아래와 같이 Fig. 4에서와 같이 상위 20개 키워드가 차지하는 80위 순위내의 연결 중심성과 매개 중심성의 순위를 비교해보았다.

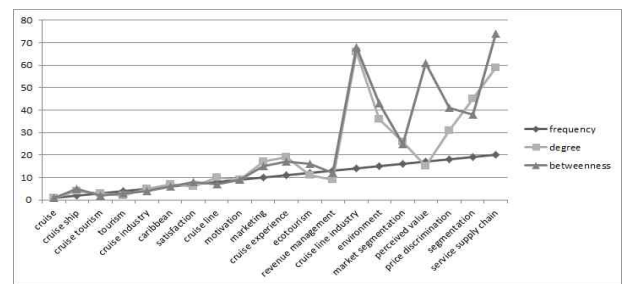


Fig. 4 Comparison of Ranking of Top 20th of Keyword Frequency with that of Degree Centrality and Betweenness Centrality

위의 Fig. 4에서 알 수 있듯이 상위 20개 중에서 14개는 연결 중심성과 매개 중심성은 빈도의 순위와 크게 다르지 않게 움직이고 있다. 따라서 본 연구에서는 주요 노드로 사용되는 키워드를 통하여 크루즈산업의 동향을 분석하는 것이 목적이므로 연결 중심성과 매개 중심성을 사용하였다.

먼저 연결 중심성은 네트워크의 연결정도에 기반하여 산출되는 지표인데 각각의 노드가 직접 연결된 이웃하는 노드가 많을수록 중심성이 높아진다(Freeman, 1979). 중심성 알고리즘의 계산공식은 아래 (2)와 같다.

$$\text{degree centrality of node} = \frac{\text{sum [weight of incident links]}}{\# \text{nodes} - 1} \quad (2)$$

연결 중심성은 노드 n이 직접 연결되어 있는 링크 수의 합을 전체노드 수에서 1을 뺀 값으로 나누어 계산한다. 연결 중심성을 통해서 산업에서 가장 중심적으로 연구되고 영향력이 가진 큰 연구주제를 파악할 수 있다. 연결 중심성은 넷마이너 프로그램에서 Analyze >> Neighbor >> Degree를 클릭하여 계산하였다.

다음으로 매개 중심성은 노드들 간의 최단 경로에 많이 포함된 노드 즉 가장 많이 거치게 되는 노드가 중심이 높게 나타난다. 매개 중심성이 높은 노드는 정보의 흐름에 대한 통제력을 가질 수 있으며 이 노드를 제거하면 키워드 네트워크의 전체 연결에 많은 영향을 주게 된다. 매개 중심성의 계산 공식은 아래 (3)과 같다.

$$c(i) = \sum_{j < k} g_{jk}(i) / g_{jk} \quad (3)$$

위의 (3)의 매개 중심성 공식에서  $g_{jk}(i)$ 는 노드 i를 통해서 지나가는 경로를,  $g_{jk}$ 는 노드 간 최단 경로를 의미한다. 프로그램에서 Analyze >> Centrality >> Betweenness >> Node를 사용하여 구할 수 있다. 매개 중심성이 높게 나타나는 키워드는 전체 키워드에서 소수에 불과하기 때문에 전체 시기별표는 제시하지는 않고 필요에 따라 내용에 추가하였다.

연결 중심성 알고리즘을 활용하여 시기별로 크루즈산업의 연구동향을 파악하기 위하여 우선 연도 기간별 연결 중심성 상위 10위의 키워드를 도출한 결과는 아래 Table 3과 같다.

Table 3 Top 10 Keywords Extracted from Degree Centrality in Each Period

Rank	1982-1999	2000-2004	2005-2009	2010-2014
1	tourism	cruise	cruise ship	cruise
2	cruise	cruise tourism	cruise	cruise tourism
3	marketing	tourism	tourism	cruise ship
4	cruise industry	satisfaction	cruise industry	tourism
5	culture	caribbean	cruise tourism	cruise industry
6	trend	cruise industry	ecotourism	motivation
7	distribution channels	pricing	perceived value	caribbean
8	virtual reality	star cruise	satisfaction	satisfaction
9	internet	p & o	customer satisfaction	revenue management
10	central reservation systems	norwegian cruise line	cruise line	cruise experience

위의 Table 3에서 보는 바와 같이 5년 주기로 각 기간별 공통적으로 연결 중심성이 높은 최상위의 키워드인 *cruise*와 *tourism* 노드를 중심으로 연도 기간별로 어떠한 변화를 가져왔는지 매개 중심성이 높은 노드들의 키워드 네트워크 시각화로 크루즈산업의 연구동향을 살펴본다. 첫 번째 기간(1982-1999)의 특이한 것은 *marketing*(0.111), *culture*(0.09), *trend*(0.08), *internet*(0.06) 등으로 크루즈산업의 태동시기의 주요한 키워드에서 우리는 2000년 전에는 크루즈선사들의 크루즈관광 시장 및 선박의 기술에 초점을 맞추고 있음을 알 수 있다. 두 번째 기간(2000-2004)의 5년간은 *satisfaction*(0.09), *Caribbean*(0.077), *pricing*(0.049), *Star Cruise*(0.042), *Norwegian Cruise Line*(0.042) 등이 나타남으로써 승객의 만족과 가격이 중요시되었고 특정 크루즈선사들의 등장과 크루즈관광지의 선호도가 앞의 2000년 이전과는 확연히 차별화된다. 세 번째 기간(2005-2009)에서는 앞의 기간들과는 달리 특이하게 나타난 *cruise ship*(0.134), *ecotourism*(0.05), *cruise line*(0.013) 등에서 크루즈선사의 크루즈선의 대형화 및 그에 따른 환경보호의 관광이 주요 주제가 되었음을 알 수 있다. 끝으로 최근 2010-2014의 5년에서는 *motivation*(0.073), *revenue management*(0.046), *cruise experience*(0.011) 등에서 크루즈관광의 동기부여와 경험 및 이익창출을 위한 고객만족 경영 기법 등을 강조하고 있다.

## 4. 연구결과

### 4.1 첫 번째 기간 (1982-1999)

2000년 이전인 1982년-1999년에 발표된 연구논문의 키워드 네트워크를 연결 중심성으로 시각화한 결과는 큰 중심축 군집 네트워크는 1개이고 선형은 7개, 파당은 4개로 나타났다. 이 기간에 나타난 연결 중심성의 상위 1위와 2위인 *tourism*(0.232)과 *cruise*(0.162)의 노드를 연결하는 매개가 무엇인지를 아래 Fig. 5에서 살펴보자.

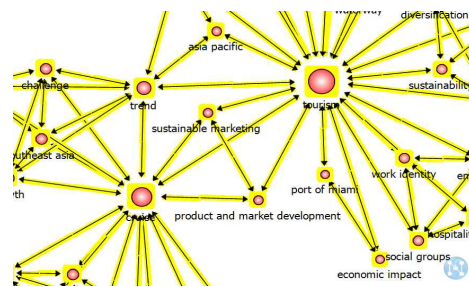


Fig. 5 Nodes between *tourism* and *cruise* between 1982 and 1999

위의 Fig. 5에서 *tourism*과 *cruise* 노드는 큰 중심축 군집 네트워크의 중심에 위치하고 두 노드가 직접 연결되어 있고 *trend*(0.08), *sustainable marketing*(0.03), *product and market development*(0.03) 등 세 개의 키워드 노드를 매개로

간접적으로도 연결된다.

#### 4.2 두 번째 기간 (2000-2004)

두 번째 기간인 2000년에서 2004년까지 5년간의 키워드 네트워크를 살펴본 결과는 아래 Fig. 6과 같이 큰 중심축 군집 네트워크는 2개이고 선형은 10개, 과당은 7개로 나타났다.

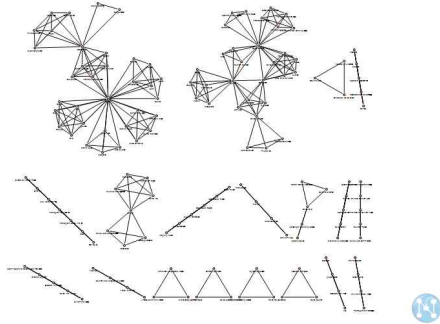


Fig. 6 Visualization of Keyword Networks between 2000 and 2004

이 기간에서는 첫 번째 기간에서 관찰된 바와는 달리 흥미롭게도 *cruise*(0.19)와 *tourism*(0.092)의 키워드 노드는 각각 1위와 3위를 차지하지만 두 개의 큰 중심축 군집 네트워크의 중심에 각각 분리되어 나타나서 두 개의 노드를 이어주는 매개는 없다.

그러면 *cruise*와 *tourism*이 각각의 큰 중심축 군집 네트워크에서 어떤 중심성이 높은 노드와 연결되어 있는지 살펴보자.

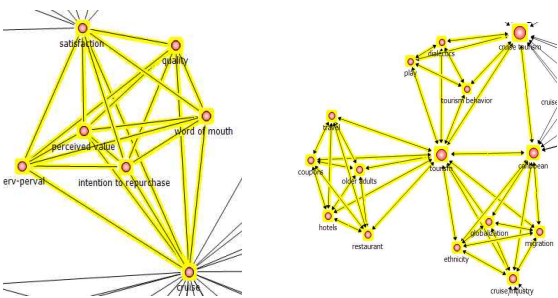


Fig. 7 Nodes between *cruise* and *satisfaction & tourism* and *cruise tourism* between 2000 and 2004

위의 Fig. 7의 좌측 네트워크 그림에서 보는 바와 같이, *cruise*는 Fig. 6의 첫 번째 큰 중심축 군집 네트워크의 중심에서 이 기간의 연결 중심성의 4위인 *satisfaction*(0.092)와 직접 연결되거나 *quality*(0.04), *perceived value*(0.04), *word of mouth*(0.03), *intention to repurchase*(0.04), *serv-perva*(0.04) 등의 4개의 노드를 매개로 간접적으로 이어준다. 그리고 위의 Fig. 7의 우측 네트워크 그림에서 보는 바와 같이 *tourism*은 두 번째 큰 중심축 군집 네트워크의 중심에서 이 기간의 연결 중심성의 2위인 *cruise tourism*(0.127)와 직접 연결되거나 5위

인 *Caribbean*(0.077), *play*(0.02), *dialects*(0.02), *tourism behavior*(0.02) 등의 노드를 매개로 간접적으로 이어준다.

#### 4.3 세 번째 기간 (2005-2009)

세 번째 기간에서는 앞선 두 기간과는 달리 흥미롭게도 아래 Fig. 8과 같이 큰 중심축 군집 네트워크가 2개, 작은 중심축 군집 네트워크가 2개, 선형이 17개 과당이 5개로 나타났다.

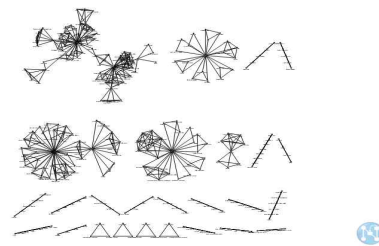


Fig. 8 Visualization of Keyword Networks between 2005 and 2009

두 개의 큰 중심축 군집 네트워크의 중심에는 왼쪽 위의 큰 중심축 군집 네트워크에서는 *cruise*(0.125)와 *tourism*(0.104)이, 왼쪽 아래에 있는 큰 중심축 군집 네트워크에서는 *cruise ship*(0.135)과 *ecotourism*(0.051)이 각각 차지하고 있다. 또한 독자적으로 새로운 네트워크를 형성하고 있는 두 개의 작은 중심축 군집 네트워크의 중심에는 오른쪽 위에서는 *cruise tourism*(0.067)이, 오른쪽 아래에는 *cruise industry*(0.088)가 있다.

세 번째 기간에서는 *cruise*와 *tourism*의 키워드 노드가 각각 2위와 3위를 차지하여 위의 Fig. 8의 좌측 상단의 한 개의 큰 중심축 군집 네트워크의 중심에 나타났다. 그러나 앞서 관찰된 바와 달리 아래의 Fig. 9에서 보는 바와 같이 특이하게도 직접 연결되지 않고 1위의 매개 중심성을 가지는 *Caribbean*(0.044) 한 개의 노드만을 매개로 간접적으로 연결된다.

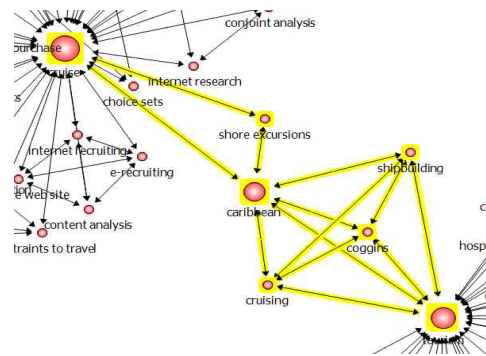


Fig. 9 Nodes between *cruise* and *tourism* between 2005 and 2009

#### 4.4 네 번째 기간 (2010-2014)

바로 앞의 세 번째 기간에서 크고 작은 중심축 군집 네트워크의 중심에 독자적으로 위치하고 있었던 *cruise-tourism*, *cruise ship-ecotourism*, *cruise tourism*과 *cruise industry*가

네 번째 기간에서는 아래 Fig. 10에서와 같이 1개의 큰 중심축 군집 네트워크 안으로 통합되어 선형이 16개, 파당이 1개만 나타났다.

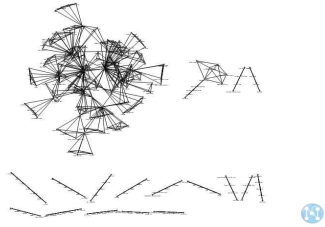


Fig. 10 Visualization of Keyword Networks between 2010 and 2014

또한 앞선 기간들과는 달리, 아래 Fig. 11에서와 같이 *cruise*(0.158)와 *tourism*(0.085)은 *cruise industry*(0.085), *Caribbean*(0.048), *cruise tourism*(0.154) 등과 함께 연결되어 있다.

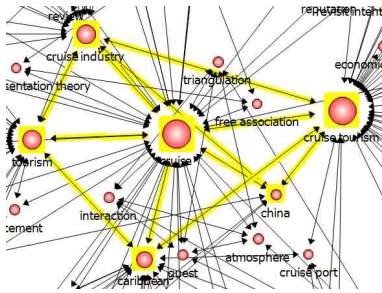


Fig. 11 Nodes between *cruise* and *tourism* between 2010 and 2014

위의 Fig. 11에서 *China*가 비교적 높은 연결 중심성 값(0.023)과 매개 중심성 값(0.014)을 가지면서 *cruise*와 *cruise tourism*을, *cruise*와 *cruise industry*를 또한 *cruise industry*와 *cruise tourism*을 연결시켜주는 매개체로 새로 등장한 것이 가장 흥미롭다. 그러므로 중국이 최근 세계 크루즈산업에서 급성장하고 있는 지역임을 고려한다면 *China*가 또 다른 어떤 노드와 연결되고 있는지 알아보는 것은 의미가 있으므로 *China*에 연결된 노드를 보여주는 아래 Fig. 12를 보자.

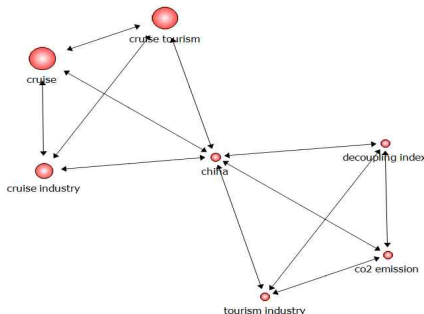


Fig. 12 Nodes Connecting *China*

위의 Fig. 12에서 보는 바와 같이 *China*는 *cruise*, *cruise tourism*, *cruise industry* 이외에도 *tourism industry*(0.011), *CO2 emission*(0.011), *decoupling index*(0.011)와도 연결되어 있다.

## 5. 결론

지금까지 키워드 네트워크 분석을 활용하여 사회연결망 프로그램인 넷마이너 4.0으로 여러 가지 중심성 측정방법 중 키워드 상위 20개의 빈도순위를 비교한 결과 빈도순위와 가장 가까운 연결 중심성과 매개 중심성을 선택하여 세계 크루즈산업의 연구동향을 분석하였다. 특히 크루즈산업의 연구동향이 연도 기간별로 어떠한 변화를 가져왔는지를 살펴보기 위해 2000년 이전과 2000년 이후로 크게 대별하여 나누고 2000년 이후에는 5년 주기로 각 기간별 공통적으로 나타나는 연결 중심성이 높은 최상위의 키워드인 *cruise*와 *tourism* 노드를 중심으로 매개 중심성이 높은 것들의 키워드 네트워크를 시각화하여 논의하였다. 그 결과는 다음과 같이 요약된다.

(i) 첫 번째 기간(1982-1999)인 2000년 이전에는 *tourism*과 *cruise* 노드는 1개의 큰 중심축 군집 네트워크의 중심에 위치하고 두 노드가 직접 연결되어 있으면서도 *trend*, *sustainable marketing*, *product and market development* 등의 세계의 키워드 노드를 매개로 간접적으로도 연결되어 있다. (ii) 두 번째 기간(2000-2004)인 2000년 이후 5년간은 *cruise*와 *tourism*의 키워드 노드는 각각 상위 1위와 3위를 차지하지만 두 개의 큰 중심축 군집 네트워크의 중심에 각각 분리되어 나타나서 두 개의 노드를 이어주는 매개는 없었다. (iii) 세 번째 기간(2005-2009)에서는 *cruise*와 *tourism*의 키워드 노드는 각각 2위와 3위를 차지하여 한 개의 큰 중심축 군집 네트워크의 중심에 나타나지만 특이하게도 직접 연결되지 않고 *Caribbean* 한 개의 노드만을 매개로 간접적으로 연결되었다. (iv) 네 번째 기간(2010-2014)에서는 세 번째 기간에 크고 작은 중심축 군집 네트워크의 중심에 독자적으로 위치하고 있던 *cruise-tourism*, *cruise ship-ecotourism*, *cruise tourism*, *cruise industry*는 1개의 큰 중심축 군집 네트워크 안으로 통합되어 나타났고 *cruise*와 *tourism*은 *cruise industry*, *Caribbean*, *cruise tourism*과 함께 연결되어 있었다. 여기서 특이하게도 최근에 세계 크루즈산업에 큰 영향을 주고 있는 *China*가 비교적 높은 연결 중심성과 매개 중심성 값을 가지면서 *cruise*와 *cruise tourism*을, *cruise*와 *cruise industry*를 또한 *cruise industry*와 *cruise tourism*을 연결시켜주는 매개체로 새로 등장하였다.

본 연구에서 사용한 키워드 빈도는 Zipf의 법칙을 따르고 노드의 연결정도는 멱함수 분포를 보여주었으므로 언어네트워크에서 하는 분석하는 방법과 동일하게 크루즈산업과 관련된 학술논문에서 제시하고 있는 키워드를 갖고 키워드 네트워크를 구축하였다. 또한 복잡계의 네트워크 분석에서 사용한 연결 중심성과 매개 중심성 분석방법으로 시대별로 나누어 시

각화하여 살펴봄으로써 세계 크루즈산업의 연구동향을 심층적으로 관찰할 수 있었다. 특히 네 번째 기간인 2010-2014에서 연결 중심성과 매개 중심성이 비교적 높은 값을 갖는 *China*의 등장과 연결노드의 키워드의 특징은 *China*가 세계 크루즈산업에서 주요한 역할을 하고 있음을 보여주므로 본 연구에서 사용한 키워드 네트워크 분석은 기간별 연구동향을 파악하는데 유용한 방법임을 보여 주었다.

본 논문의 연구결과가 가지는 산업적 의의는 지금까지 크루즈산업에서 중요시 되고 있던 주제가 무엇인지 시간적으로 파악하여 향후 중요하게 다루어질 산업분야의 동향 예측의 기초 자료로 활용할 수 있다는 점이다. 또한 학문적 의의는 키워드 네트워크 분석을 통하여 확인된 연결중심성이 높게 나온 키워드들은 크루즈산업을 이해하려는 사람들에게 최신의 연구동향 및 연구주제를 파악할 수 있는데 도움을 줄 수 있다는 것이다. 한편 본 연구의 한계점은 국제저널에 게재된 240개의 적은 수의 논문을 대상으로 연구한 것이라 본다. 따라서 더 많은 수의 논문을 포함시킨다면 보다 포괄적인 크루즈산업의 모습을 파악할 수 있었을 것으로 본다. 추후 연구로 좀 더 많은 수의 논문을 대상으로 저자를 매개로 한 키워드 네트워크 분석으로 크루즈산업을 연구하는 연구자 맵을 구축하는 것도 세계 크루즈산업을 연구하고 이해하는 데 큰 도움이 될 것이라 본다.

## References

- [1] Borgatti, S. P., Mehra, A., Brass, D. J. and Labianca, G.(2009), "Network Analysis in the Social Sciences", *Science*, Vol. 323, No. 5916, pp. 892-895.
- [2] Chaurasia, N. and Tiwari, A.(2014), "On the Use of Brokerage Approach to Discover Influencing Nodes in Terrorist Networks", *Intelligent Systems Reference Library*, Vol. 65, pp. 271-295.
- [3] Cho, S. C., Won, Y. K., and Kim, J. H.(2012), "A Network Optimization Model for Strategic Itinerary Planning of Cruise Fleet (Korean version)", *Journal of Navigation and Port Research*, Vol. 36, No. 1, pp. 51-58.
- [4] Cyram(2013), *Netminer 4.0*. Seoul: Cyram Co. Ltd.
- [5] Freeman, L. C.(1979), "Centrality in Social Networks: Conceptual Clarification", *Social Networks*, Vol. 1, pp. 215-239.
- [6] Jang, H. L., Kang, G. W. and Lee, E. J.(2012), "Analysis of Research Subject Network in the Field of Oncogene (Korean version)", *Journal of Korea Technology Innovation Society*, Vol. 15, No. 2, pp. 369-399.
- [7] Jeong, C. W. and Kim, J. J.(2012), "Analysis of Trend in Construction Using Textmining method (Korean version)", *Journal of the Korean Digital Architecture and Interior Association*, Vol. 12, No. 2, pp. 53-60.
- [8] Kho J. C., Cho, K. T. and Cho, Y. H.(2013), "A Study on Research Trend in Management of Technology Using Keywords Network Analysis (Korean version)", *Journal of Intelligent Information Systems*, Vol. 19, No. 2, pp. 101-123.
- [9] Kwon, S. Y.(2014), "A Study on the Application to Network Analysis on the Importance of Author Keyword based on the Position of Keyword (Korean version)", *Journal of the Korean Society for Information Management*, Vol. 31, No. 2, pp. 121-142.
- [10] Kwon, K. S., Yi, J. Y., Lee, J. Y.(2014), Chae, S. W. and Han, D. S.(2014). "A Social Network Analysis on the Research Trend of Korean Medicine (Korean version)", *Journal of Korea Technology Innovation Society*, Vol. 17, No. 2, pp. 1-21.
- [11] Lee, S. S.(2012). *Network Analysis Methods (Korean version)*. Seoul: Nonhyung.
- [12] Lee, J. and Kang, J. H.(2011), "Mapping Knowledge Domains: The Structure of Knowledge Production and the Status of Theory in Korean Sociology 2004-2010 (Korean version)", *Korean Society for Social Theory*, No. 19, pp. 89-144.
- [13] Liang, W., Shi, Y., Tse, C. K., Liu, J., Wang, Y. and Cui, X.(2009), "Comparison of Co-occurrence Networks of the Chinese and English Languages", *Physica A*, 388, pp. 4901-4909.
- [14] Newman, M. E. J.(2005), "Power Laws, Pareto Distributions and Zipf's Law", *Contemporary Physics*, Vol. 46, No. 5, pp. 323 - 351.
- [15] Park, J. Y., Fu, Z. J. and Chiu, D. M.(2012) "Networking, Clustering and Brokering Keywords in the Computer Science Research", In *Proceedings of Digital Information Management*, pp. 110-115.
- [16] Ryu, S. I. and Hyun, T. S.(2013), "Analysis of Research Trends in the Welfare Field of Public Administration (Korean version)", *The Korea Contents Association*, Vol. 13, No. 2, pp. 254-262.
- [17] Yang, S. D.(2013), "Semantic Network Analysis of Research Trend Related to Private Security (Korean version)", *The Korea Contents Association*, Vol. 13, No. 11, pp. 894-901.
- [18] Zipf, G. K.(1949), *Human Behavior and the Principle of Least Effort*, Oxford: Addison-Wesley Press.

Received 18 November 2014

Revised 16 December 2014

Accepted 17 December 2014