

# 소셜네트워크 분석을 통한 유라시아 대륙철도 관련 키워드 변화 연구

## A Study on Key Words Change of TER(Trans Eurasian Railway) by SNA(Social Network Analysis)

진장원\*, 신동희†

JangWon Jin\*, DongHee Shin†

**초 록** 본 논문은 정치적 환경에 민감한 특정키워드(유라시아 대륙철도, 남북철도, 평화철도)의 변화 과정에 대해 SNS(Social Network Service)에서 수집된 빅데이터를 이용하여 분석한 것이다. 수집된 빅데이터 채널은 네이버, 다음의 블로그 및 카페이며, 수집기간은 남북 관계가 경색되어 있었던 박근혜 전 대통령 재임기간(수집기간 1)과 남북관계가 풀리기 시작한 현 정부 기간(수집기간 2)이다. 분석방법은 양 기간 동안에 있었던 유라시아 대륙철도에 대한 SNS 사용자의 생각과 의견을 Python3.6을 이용하여 수집, 정제하고 SNA(Social Network Analysis)를 실시하였다. 그 결과 수집기간 1에서는 특정 키워드에 대한 단순개체명 단어빈도(지리적 명칭)가 다수 나타났으나, 수집기간 2에서는 특정 명사(북한, 미국, 북핵 등)와 구체적인 명사들(연결, 추진, 개발 등)이 집중적으로 나타났다. 이러한 결과를 통해 수집기간1보다는 수집기간2가 특정키워드에 대한 SNS사용자들의 의견과 생각들이 더 많이 내포되어 있으며, 동시에 네티즌들은 유라시아 대륙철도 정책이 북한의 핵실험, 미사일 발사실험 등 국제정세에 영향을 많이 받고 있다고 인식하고 있음을 알 수 있었다. 이러한 빅데이터 분석은 향후 정부가 본 정책을 추진하는데 있어서 국민들을 설득하는데 효과적인 방향을 제시 할 수 있을 것으로 예상 된다.

**주요어** : 빅데이터, SNS(Social Network Service), SNA(Social Network Analysis), Python3.6, R 3.4.3

### 1. 서 론

유라시아 대륙철도 연결은 단순히 교통인프라만을 구축하는 사업이 아니라 남북간의 긴장을 완화해주고 통일을 준비시켜주는 정치적 이슈와도 맥을 같이 하는 중요한 의제이다. 하지만 2008년 경의선 철길이 끊긴 후 9년 동안 유라시아 대륙철도 연결 사업은 실제적으로 정체상태였다가 현 정부 들어와서 북방경제위원회 등을 통해 다시 부활하고 있다. 한편 최근 빅데이터 분석 결과를 활용하여 정부정책과 기업경영 혁신과 관련한 뉴스 보도가 많아지면서 빅데이터에 대한 사회적 관심이 급증하고 있다.

한 예로 미국 대선(2016.11) 당시 대다수 미국언론과 여론조사에서는 힐러리 클린턴 후보가 승리할 것으로 예상했으나 결과는 빅데이터 분석이 예측했던 대로 트럼프 후보가 당선 되었다. 이로 인해, 빅데이터에 대해 국내·외에서 많은 관심이 집중되기 시작했다.

따라서 본 논문은 정치적 환경에 민감한 특정키워드(유라시아 대륙철도, 남북철도, 평화철도)의 변화 과정에 대한 분석을 위하여 SNS에 나타난 사용자들의 의견, 생각들을 빅데이터 분석을 통해 유라시아 대륙철도 연결 정책 추진 시 나가야할 방향을 제시 하고자 하였다. 연구방법은 첫째, Python3.6과 R을 이용하여 빅데이터의 수집 및 정제를 하고 둘째, Ucinet6을 이용하여 빅데이터의 SNA를 시도하였다.

\* 주저자: 한국교통대학교 교통대학원 교수

† 교신저자: 한국교통대학교 교통정책시스템공학과 박사과정([kant114@naver.com](mailto:kant114@naver.com))



Fig. 1 Collection 1(L), Collection 2(R) Text Mining

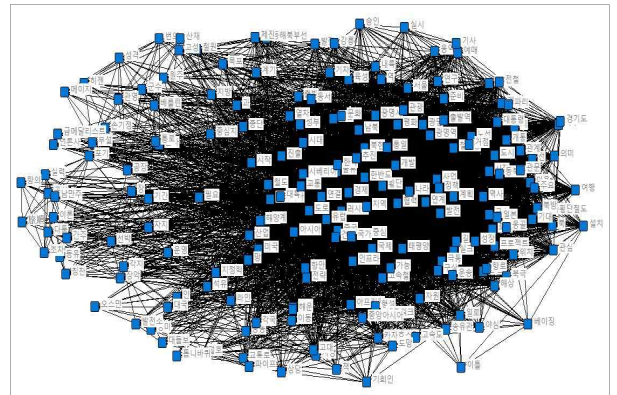
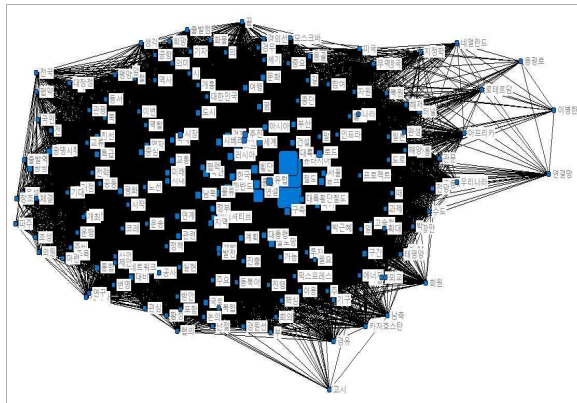


Fig. 2 Collection 1(L), Collection 2(R) SNA(Social Network Analysis)

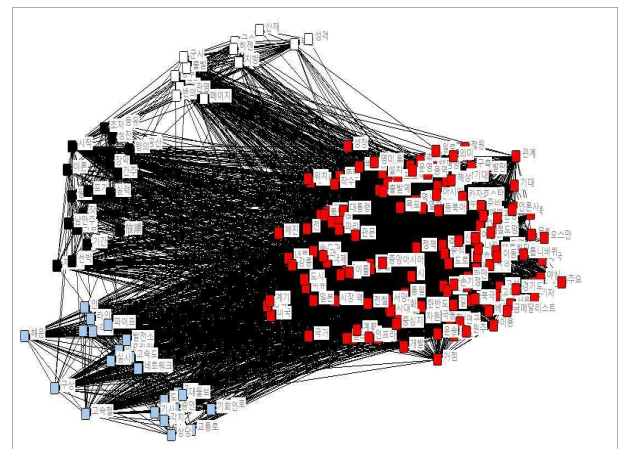
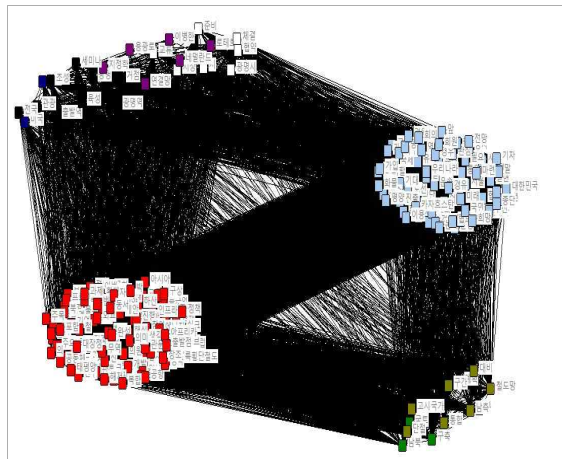


Fig. 3 Collection 1(L), Collection 2(R) CONCOR(CONvergence of iterated CORrelations)

## 2. 본 론

### 2.1 연구방법 및 조사설계

빅데이터란 아날로그 환경에서의 데이터에 비해 생성되는 주기가 짧고 수치 데이터와 텍스트, 이미지를 포함하는 방대한 규모의 데이터를 의미한다[1]. 빅데이터 분석에는

데이터마이닝, 텍스트마이닝, 소셜 네트워크 분석, 클러스터 분석 등이 있다[2]. 본 논문에서의 빅데이터 분석은 첫째, 수집은 최근에 가장 활발히 이용하고 있는 Python3.6을 이용하여 특정키워드에 대한 빅데이터를 수집하였다. 둘째, 수집된 빅데이터의 정제(명사추출, 특수문자, 동사, 형용사 제외)는 R

을 이용하여, 텍스트마이닝분석을 하였다. 셋째, 소셜네트워크(SNA) 및 클러스터분석(CONCOR)은 Ucinet6을 이용하였다. 해당 프로그램들은 무료프로그램이라는 강점을 가지고 있어, 국내·외 활발히 이용되고 있다. 특정키워드에 대한 빅데이터 수집 채널은 국내 포털사이트(naver, daum)의 블로그, 카페로 하였다(쿼리 1000으로 제한, facebook, twitter는 생성주기가 6개월로 제외 함). 수집된 빅데이터는 Table1과 같다.

**Table 1** Collection of Big data

Channel	Term	
	Collection 1	Collection 2
naver	3,375cases	2,240cases
daum	2,950cases	1,720cases

수집된 빅데이터의 건수를 살펴보면 네이버에서 빅데이터가 더 많이 수집되었다. 그 이유는 네이버가 다음보다 SNS이용자가 많다는 것을 나타낸다.

### 2.1.1 빅데이터 정제 및 SNA

수집된 빅데이터 정제는 R을 이용하여 텍스트마이닝을 실시하였다. 텍스트마이닝 분석 결과는 Fig1.과 같다. 특정키워드에 대한 단어 출현빈도는 수집기간 1.에서는 철도(2791건), 유라시아(2492건), 대륙(2054건)으로 나타났고, 수집기간 2.에서는 중국(270건), 유라시아(150건), 러시아(125건), 광명시(95건)순으로 나타났다. 소셜 네트워크분석(SNA)은 개인들 간의 사회적 관계를 구조적으로 분석하여 내재된 관계를 파악하는 방법, 기술이다[3]. SNA분석 프로그램은 Ucinet6 Netdraw를 이용하였고, 전체 네트워크는 Fig2 과 같다. SNA 분석 후 출현단어의 구조적 등위성 분석을 알아보기 위하여 CONCOR분석을 실시하였다. CONCOR분석은 텍스트가 전달하고자 하는 유의미한 키워드를 분석하고 키워드 간의 네트워크를 통하여 상호 관계성을 파악한다[4].

## 3. 결론

본 연구에서 특정키워드(유라시아, 대륙철도(TER), 남북철도, 평화철도)에 대한 빅데이터 수집 후 텍스트마이닝 결과 수집기간1. 과 수집기간 2. 에서는 특정키워드에 대한 출현빈도 단어는 수집기간1. 에서는 다양한 특정키워드에 대한 단어들 수집 되었고, 수집기간 2.에서는 북한, 미국 등 특정국가 명사들이 수집되었다. 이러한 결과는 수집기간 2. 기간에 북한의 핵실험과 미사일 시험에 대한 특정키워드가 영향을 받은 것으로 예상 할 수 있다. 수집기간 1.에서의 KTX광명역이 수집기간2.에서는 광명시가 더 많은 출현빈도를 나타냈다. 이러한 이유는 광명시가 TER출발역으로 KTX광명역을 중점 홍보한 것으로 인해 SNS사용자들에게 많이 노출된 결과임을 유추해볼 수 있다. Fig2는 전체 네트워크 분석이며, 수집기간1.의 전체네트워크에서 중심성을 나타내는 단어들은 아시아, 유럽, 한국, 중국, 시베리아, 등 지명 단어명이 중심에 내포되어있고, 수집기간2.에서 중심성을 나타내는 단어들은 특정키워드에 대한 구체적인 단어들 즉, 연결, 교통, 추진, 개발 등이 나타났다. Fig3의 CONCOR분석 결과 수집기간1.에서 보다 수집기간2.에서 특정 키워드에 대한 기대와 구체적인 클러스터 개체가 더 많이 도출되었다. 이러한 결과를 통해 SNS사용자들이 수집기간1.보다 수집기간2.에서 TER도에 대해 더 많이 희망과 기대감을 갖고 있는 것으로 분석되었다. 물론 이 연구는 SNS분석으로 한정되지만, 이러한 시도는 향후 정부가 TER정책을 추진하는데 있어서 국민들을 설득하는데 효과적인 방향을 제시 할 수 있을 것으로 예상 된다.

## 참고문헌

- [1] 한국정보화진흥원(2012) 미래 IT 기술발전 방향 및 전망, 정책연구, 17, pp 1-61.
- [2],[3] R을 이용한 빅데이터 분석(2016, 임동훈, 자유아카데미)
- [4] Boragtti,S.P.,Everett,M.G.,& Freeman,L.C.(2002). UCINET 6 for windows: Software for social network analysis(Version 6). Harvard,ma:analytic technologies.