

텍스트마이닝 기법을 활용한 부산시 지역혁신정책 동향분석*

박 주 섭**
홍 순 구***

국문요약

현대 사회에서는 다양한 사회문제가 발생하고 있고 지역정부에서는 이러한 문제를 해결하기 위해 많은 노력을 기울이고 있다. 해결 방안을 모색하기 위해서는 먼저 당면한 지역사회정책 특히 지역 경제성장 및 발전과 연관된 지역혁신정책의 동향을 살펴볼 필요가 있다. 본 연구의 목적은 지역혁신정책의 동향을 분석하고 다른 지역혁신정책과 연관성이 높으면서 영향력이 큰 지역혁신정책을 살펴보고 그 시사점을 도출하는 데 있다. 이를 위해 빅데이터를 활용한 데이터 분석 기법중 비정형 데이터 분석의 하나인 텍스트마이닝 기법을 활용하여 2002년부터 2015년까지 부산 시정신문인 "Buvi News" 기사를 분석하였다. 분석 결과 지역혁신에 있어서 인프라 사업의 이슈가 상대적으로 감소하는 경향을 보이고 있으며, 엔터테인먼트산업의 다양한 사업이 증가하는 경향을 보이고 있다. 또한 동·서 균형개발 사업과 해양관광 사업은 민선 5기 이후 활발히 논의되고 있다. 본 연구는 지역혁신과 관련된 핵심문제 도출을 위해 텍스트마이닝 분석 방법을 제시하였고, 후속 연구를 유발하였다는 학문적 기여도를 가진다. 또한 실제 지방정부 및 행정기관에서 정책 수립을 위한 자료로 활용할 수 있다는 실무적인 기여도가 있다.

주제어: 지역혁신정책, 네트워크 분석, 텍스트마이닝, 중심성 분석, 동향분석

I. 서론

지방은 수도권에 비해 정치, 경제, 문화, 교육 등 많은 분야에서 소외되어 왔다. 참여정부에서는 국가균형발전을 국정과제로 채택하고 지역불균형문제를 해소하기 위해 많은 노력을 기울였으나 수도권과 비수도권의 격차는 줄어들지 않고 있다. 예를 들어 오상영(2015)은 사람과 기업의 수도권 집중 현상의 가속화로 지방의 산업, 인재, 인구 공동화 문제는 재앙 수준으로 치닫게 될 것이라고 하였다.

* 이 논문은 2015년 대한민국 교육부와 한국연구재단의 지원을 받아 수행된 연구임
(NRF-2015S1A3A2046781)

** 제 1저자

*** 교신저자

지방정부들은 지역시민들의 삶의 질을 높이기 위해 지역혁신을 강조하여 왔지만 그 동안 혁신에 관한 논의들은 주로 지역혁신체제 구축에 집중되어 왔으며, 지역혁신 문제를 해결하기 위해 필요한 작업인 지역혁신에 대한 문제를 정의하고 범위를 설정하는 것에 대한 논의는 소홀히 해왔다. 이로 인해 지역혁신 정책에 대한 연구가 초기에 머물러 있는 상황이라 국내에선 사례분석도 많지 않다. 지역혁신 문제를 정의하고 범위를 설정하기 위해선 실제 지역혁신 정책의 이슈가 어떻게 변화해 왔는 지에 대한 동향 분석이 반드시 필요하다. 이를 통해서 지역혁신 정책의 흐름을 파악할 수 있고 후속 연구를 유발할 수 있으며 또한 실제 지방정부 및 행정기관에서 정책 수립을 위한 자료로 활용할 수 있다. 소셜 네트워크 등 IT 발전으로 인해 정형화된 데이터 뿐만아니라 비정형 데이터까지도 처리(박두순, 2014)할 수 있게 됨에 따라 텍스트마이닝 기법이 사회학, 행정학 등에서 동향분석 연구에 활발히 활용되고 있다. 이러한 연구는 크게 연구 동향분석(김하진 외, 2014; 임시영 외, 2014; 최영출 외, 2011), 기술개발 동향분석등 기술분야에서의 연구(감주식 외, 2013; 이인석 외, 2015) 및 사회위험 이슈 도출 등 사회분야(구주나, 2014; 양현수, 2012; 원진영 외, 2014; 정다운 외, 2015; Mohamed M. Mostafa, 2013)의 연구로 구분될 수 있다.

이 연구의 목적은 다음과 같다. 첫째, 텍스트마이닝 기법을 활용하여 부산시의 지역혁신정책 동향을 분석하는 것이다. 둘째, 다양하고도 많은 지역혁신문제를 한번에 모두 풀 수는 없기 때문에 다른 지역혁신문제와 연관성이 높고 가장 영향력이 큰 지역혁신정책을 도출할 필요가 있다. 이를 위해 부산광역시에서 시정 정보를 전달하는 신문인 "Buvi News" 신문기사를 분석대상으로 선정한다.

분석시기는 부산시장의 임기에 맞춰 민선 3기(2002.7.1. ~ 2006.6.30.), 민선 4기(2006.7.1. ~ 2010.6.30.), 민선 5기(2010.7.1. ~ 2014.6.30.), 민선 6기(2014.7.1. ~ 2015.10.31.)로 구분하였다. 분석 방법으로는 지역혁신 핵심어 빈도분석, 네트워크 분석, 연결 중심성 분석, 매개 중심성 분석을 실시한다.

본 논문의 구성은 다음과 같다. 2장에서는 혁신 및 지역혁신을 정의하고 텍스트마이닝을 활용한 동향분석들을 요약하며, 3장에서는 핵심어 빈도 분석, 네트워크 분석, 연결 중심성 분석, 매개 중심성 분석을 실시한다. 4장에서는 결론을 제시한다.

II. 이론적 배경

1. 지역혁신의 개념 및 정의

혁신은 사회과학 및 자연과학을 비롯한 거의 모든 다양한 학문 분야에서 다루어지고 있다. 특히 빠른 변화가 진행되고 있는 정보화 사회에서의 혁신은 다양한 의미를 내포하고 있고(이승현, 2012) 여러 영역에서 활용되는 만큼 그 개념도 다양하게 정의되고 있다.

혁신에 대한 초기 논의로 기술혁신 측면에서 Schumpeter(1934)는 신제품, 새로운 생산방법, 새

로운 판로의 개척, 원료·반제품의 새 공급원 및 새로운 사업조직·생산조직·산업조직 등 다섯 가지 생산요소의 새로운 결합을 수행하는 것이라 하였다(조운애 외, 2005, 재인용).

김현민(2001)은 “혁신은 기업에 적용되던 것을 공공부문의 비효율성을 개선시키기 위해 기존 서비스와 고객만족 향상과 정부 능률성을 제고하기 위한 관리·수행방식의 새로운 변화”라고 정의했다. 김정홍(2005)은 “혁신은 협의와 광의로 나누어 정의되며 협의의 개념은 신제품, 새로운 서비스, 새로운 생산방법이 상업적 목적으로 처음 도입되는 것을 의미하며 광의의 개념은 신제품, 새로운 서비스 및 새로운 방식을 만들어 내는 데 기여하는 모든 활동을 의미한다”라고 했다.

결론적으로 혁신은 ‘기존과 다르게 새롭게 변화시키는 것’으로 정의할 수 있으며 그 변화는 혁신의 대상이나 범주에 따라 다양한 혁신이 나오고, 새로운 시장 및 새로운 비즈니스를 창출함으로써 경제가 성장하고 발전하는 데 기여한다.

혁신의 개념은 대상이나 범위가 넓은 광의의 개념(이승현, 2012)인 반면에 지역혁신은 지역이라는 행정구역내에서 경제성장 및 발전에 원인이 되는 혁신의 일부분이다. 지역의 사전적 의미는 일정한 지표면의 범위를 가르키는 지방 또는 지구 등과 거의 동의어로 쓰이는 데, 학술상으로는 일정한 목적과 방법에 의하여 구획되어진 특색있는 지구를 말한다(두산백과). 본 논문에서의 지역은 지방자치단체에 일정범위의 자치권을 부여하고 관리시키는 행정구역으로 정의 한다.

지역혁신의 결과로 새로운 시장 및 새로운 비즈니스를 창출하게 되는 데 이로 인해 지역내 기득권을 지키려는 기존 시장 및 기존 비즈니스 이해당사자들의 참여로 갈등을 초래할 수 있다. 결국 기존 이해당사자와 신규 이해당사자의 참여정도와 갈등정도에 따라 지역혁신정책을 분류할 수 있다.

본 논문에서는 지역혁신정책을 지역의 경제성장 및 발전에 기여하는 혁신적인 사업을 추진하는 것으로 이해당사자들의 참여범위가 큰 사업으로 정의한다.

2. 텍스트 마이닝을 활용한 동향분석 선행연구

텍스트 마이닝(text mining) 또는 텍스트 기반의 지식 발견(knowledge discovery from text)은 Feldman et al.(1995)에 의해 최초 언급되어졌고, 비정형 텍스트 기반 데이터베이스에서 자연어처리를 기반으로 새로운 정보를 추출하는 기술이다(Hotho et al. 2005).

텍스트 마이닝을 활용한 동향분석 선행연구들을 살펴보면 학술분야, 기업·기술 분야, 사회분야로 세 가지로 분류될 수 있다.

첫째로, 학술분야에서 보면 최영출 외(2011)는 텍스트에 출현하는 단어와 단어사이의 관계를 링크로 표시함으로써 구축되는 네트워크를 통해 현상을 해석하는 네트워크 텍스트 분석방법을 행정학 연구 경향을 분석하는 데 적용하였다. 이 연구는 단순한 주제별 분류방식이 아니라 단어간 관계성에 초점을 두고 핵심적인 연계개념들을 도출하였다는 특징이 있다. 임시영 외(2014)는 공간정보 분야의 연구동향을 파악하기위해 국가연구개발보고서와 논문을 추출하여 키워드에 대한 전처리를 수행한 후 분야별로 정리하였다. 정리된 키워드들을 통해 보고서 및 논문에서 키워드의 시기

별 출현 빈도 및 변화를 살펴보고 이를 통해 공간정보 분야의 연구동향을 확인하였다. 김하진 외(2014)는 국내외 정보학 분야 학회지 동향을 파악하기 위하여 1990년부터 2013년까지 국내 정보관리학회지와 국외 JASIST의 논문 제목과 초록을 대상으로 텍스트 마이닝 기법을 통한 명사, 명사구 동시출현 분석을 수행하였다. 전 기간을 5구간으로 나누고 전체적인 동향을 살펴보기 위해 고빈도 동시출현 단어를 분석하였으며, 세부 주제를 파악하기 위해 상위 키워드와 동시에 출현하는 단어를 분석하였다.

둘째로, 기업·기술 분야에서 보면 감주식 외(2013)는 특허를 통한 기술개발 동향분석을 위해 IPC 코드에 기반한 기술분류를 이용하여 바이오분야의 국내외 특허를 수집하고 텍스트마이닝 기법을 활용하여 유망기술 분야를 도출했다. 이인석 외(2015)는 제조시스템 분야에서 출간되는 논문(총 2,688편)을 대상으로 제조시스템 분야의 연구동향을 파악하였다. 대응분석(Correspondence Analysis, CA) 방법을 이용하여 제조시스템 분야의 연구동향 및 시각화를 했으며 행(Row)점 & 열(Colum)점 분석을 통해 실험결과를 해석했다.

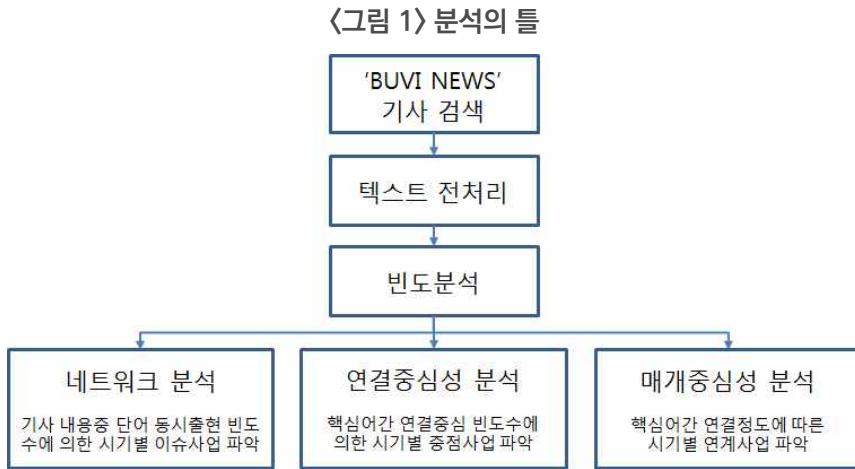
셋째로, 사회 분야에서 보면 양현수(2012)는 1990년대, 2000년대, 2010년대 초의 청소년 범죄 및 비행의 유형 변화를 알아보기 위해 텍스트 마이닝을 활용하여 인터넷 신문기사의 텍스트 데이터를 분석하였다. 분석을 위해서 인터넷 포털사이트 네이버 뉴스검색과 한국언론재단 인터넷 기사 검색 사이트 카인즈에서 청소년 범죄 기사 총 1,499건의 인터넷 기사를 수집하였다. 기사의 텍스트를 Microsoft Excel 2007 프로그램에 입력하여 데이터베이스화 한 후 SAS Enterprise Miner Client 6.2를 사용하여 분석하였다. 텍스트 데이터를 분석하기에 앞서 사전을 제작하고, 띄어쓰기 작업의 전처리 과정을 실시한 후 텍스트 마이닝을 통해 분석했다.

원진영 외(2014)는 뉴스를 대상으로 텍스트마이닝을 활용하여 토픽을 구성하는 핵심 키워드의 빈도, 문서수, 토픽 가중치 정보와 원문 검토를 통해 토픽을 도출했다. 구주나(2014)는 한국어 텍스트 마이닝의 과정을 정리하고 단계별 특징을 고찰하며 텍스트 자료를 포함하는 실제 빅데이터인 2011 경제총조사의 한식 음식점업 사업체 자료에 적용하여 분석과정에서 발생하는 통계 및 기술적 문제점들을 정리하고 해결 방법을 제시하였다. 특히 설문조사의 주관식 문항에 대한 응답 형태로 기록된 텍스트 자료에 한국어 텍스트 마이닝을 적용할 경우 발생하는 문제점들을 해결하기 위해, 기존 범용 사전이 아닌 특정 자료에 맞는 피드백 기반 사용자 사전을 구축하는 것을 제안하였다. 자료 분석을 위한 소프트웨어로는 오픈 소스 플랫폼인 R을 이용하였으며 텍스트 자료에 대한 마이닝 결과를 수치 자료로 구성된 구조화된 데이터베이스와 통합하고 군집분석 및 시각화 분석 등의 데이터 마이닝 방법들을 적용하여 전국 한식 음식점업 사업체에서 취급하는 대표 메뉴의 현황과 특성을 다각도로 탐색하였다. 정다운 외(2015)는 신문기사 중 U-City 서비스 관련 기사를 대상으로 시사분석을 통해 서비스 이슈와 동향을 분석하여 시민체감형 서비스 추진방향을 모색했다. 분석을 위해 2009년부터 2014년까지 수집된 723개 신문기사는 전처리과정을 거쳐 키워드 빈도분석을 실시했으며, 빈도분석 결과를 바탕으로 네트워크 분석과 정보계측 기법을 활용했다. Mostafa(2013)는 트윗 내용 중에 유명브랜드에 대한 감성 분석을 통해 소비자들의 심리를 분석하였다.

Ⅲ. 연구방법론

1. 분석의 틀

본 연구의 목적을 달성하기 위해 2장에서 살펴본 문헌연구와 빈도분석을 바탕으로 민선시기별 핵심사업을 살펴보았다. 연구방법으로는 네트워크 분석으로 민선시기별 이슈 핵심 사업을 파악하였고, 연결 중심성 분석으로 민선시기별 중심성 사업을, 매개 중심성 분석으로 민선시기별 연계성 사업을 각각 파악하였다. 분석의 틀을 도식화하면 다음 <그림 1>과 같다.



2. 네트워크 분석 개요

지역혁신문제를 도출하기 위해 상업 신문을 제외시키고 부산 시정 신문인 “Buvi News” 기사를 분석대상으로 선정했다. 주제어인 “지역혁신”을 키워드로 2002년 이후 기사를 추출해서 빈도분석을 통해서 1차로 연관 키워드를 검색한 후 관련문헌 검토 및 혁신 전문가 2명과 정책 전문가 1명이 브레인스토밍을 통해 ‘부산발전’, ‘경제성장’, ‘경제발전’의 연관키워드를 선정했다. ‘부산발전’, ‘경제성장’, ‘경제발전’을 키워드로 민선 3기(2002.7.1. ~ 2006.6.30.) 210개 기사와 민선 4기(2006.7.1. ~ 2010.6.30.) 266개 기사와 민선 5기(2010.7.1. ~ 2014.6.30.) 354개 기사와 민선 6기(2014.7.1. ~ 2015.10.30.) 66개 기사를 대상으로 분석하였다. 핵심어의 빈도분석은 한국어 텍스트를 분석할 수 있는 KrKwic 프로그램을 사용하였다. KrKwic은 어떤 특정한 단어들의 출현빈도 뿐만 아니라 동시 출현 단어들을 나타내 주는 특징을 가지고 있다(박한우외, 2004). 핵심어 빈도 분석 결과를 바탕으로 NetMiner 프로그램을 사용하여 Network 분석 및 중심성 분석, 매개 중심성 분석을 실시하였다. 네트워크의 구성요소는 노드와 링크가 되며, 노드(node)는 상호 고유한 속성을 가지는 행위자(actor)를 나타내며, 링크(link)는 노드들간의 연결 관계를 나타낸다(이수상, 2012).

3. 핵심어 빈도 분석

기사를 민선 3기(2002.7.1. ~ 2006.6.30.), 민선 4기(2006.7.1. ~ 2010.6.30.), 민선 5기(2010.7.1. ~ 2014.6.30.), 민선 6기(2014.7.1. ~ 2015.10.31.)로 구분하여 빈도 분석한 결과는 <표 1>과 같다. 빈도 분석 전에 지역혁신과 관련성이 있는 단어들만 골라낸 후 조사부분을 빼고 명사화시키는 1차 작업과 명사화된 단어들의 상대비교 수준을 맞추기 위한 2차 단어 추출 작업을 실시했다. 1차 작업과 2차 작업은 혁신 전문가 2명과 정책 전문가 1명이 브레인스토밍을 통해 민선 3기 15개, 민선 4기 16개, 민선 5기 18개, 민선 6기 18개의 주요 핵심어를 선정했다.

<표 1> 부산시장 재임기간별 지역혁신 주요 핵심어

구분	민선 3기 (2002.7.1 ~ 2006.6.30)			민선 4기 (2006.7.1 ~ 2010.6.30)			민선 5기 (2010.7.1 ~ 2014.6.30)			민선 6기 (2014.7.1 ~ 2015.10.31)		
	핵심어	빈도	비율 (%)	핵심어	빈도	비율 (%)	핵심어	빈도	비율 (%)	핵심어	빈도	비율 (%)
1	APEC	84	30	동남권신공항	34	19	일자리창출	47	13	일자리창출	30	20
2	부산아시아게임	44	16	북항재개발	24	13	가덕도신공항	38	10	공공기관이전	14	9
3	부산신항	31	11	부산신항	19	11	김해공항	33	9	의료관광	11	7
4	부산국제영화제	24	8	김해공항	14	8	부산시민공원	29	8	부산국제영화제	11	7
5	백스코	18	6	가덕도신공항	12	7	광안대교	25	7	부산공동어시장	9	6
6	공공기관이전	11	4	센텀시티	10	6	북항재개발	23	6	도시철도	8	5
7	명지대교	11	4	한일해저터널	10	6	동남권신공항	22	6	구덕운동장	7	5
8	부산항만공사	11	4	자갈치시장	8	4	부산국제영화제	21	6	복합리조트	7	5
9	경부고속철도	9	3	부산국제영화제	7	4	동부산관광단지	19	5	불꽃축제	7	5
10	북항재개발	8	3	시내버스 준공영제	7	4	부산신항	19	5	게임산업	6	4
11	세계한상대회	8	3	강서수변공간	6	3	크루즈	15	4	북항재개발	6	4
12	김해공항	7	2	부산항만공사	6	3	산복도로	13	4	중남미시장진출	6	4
13	센텀시티	7	2	일자리창출	6	3	컨벤션	13	4	ITU전권회의	5	3
14	컨벤션	5	2	부산구치소	5	3	지스타	12	3	가덕도신공항	5	3
15	시민스포츠타운	5	2	부산시민공원	5	3	MICE산업	10	3	백스코	5	3
16				산복도로	5	3	마을기업	10	3	MICE산업	5	3
17							에코델타시티	10	3	부산시민공원	5	3
18							감천문화마을	9	2	부산국제 금융센터	5	3

분석 결과, 민선 3기(2002.7.1. ~ 2006.6.30.)에는 APEC의 빈도 수가 84회로 가장 높은 것으로 나타났다 다음으로 부산아시아게임, 부산신항, 부산국제영화제, 백스코, 공공기관이전, 명지대교, 부산항만공사출범 순으로 나타났다. 단어의 출현빈도가 높다는 것은 2005년 APEC과 2002년 부산아시아게임의 성공적 개최로 인하여 지역기반시설이 확충하는 계기가 되었으며, 세계도시로 도약하기 위한 부산의 분위기를 반영했던 시기라는 것을 알 수 있다. 민선 4기(2006.7.1. ~ 2010.6.30.)에는 동남권신공항과 북항재개발, 부산신항의 빈도 수가 가장 높았으며, 민선 5기(2010.7.1. ~ 2014.6.30.)에는 일자리창출과 가덕도신공항, 김해공항의 빈도 수가 가장 높았다. 민선 6기

(2014.7.1. ~ 2015.10.31.)에는 일자리창출이 30회로 가장 많이 나타났다. 다음으로 공공기관이전, 의료관광, 부산국제영화제 등의 핵심어 빈도가 높게 나타났다. 특히 민선 5기와 6기에 일자리창출의 빈도수가 가장 많이 나타난 것은 경제의 어려운 실상을 반영하고 있는 것으로 보인다. 그리고 민선 5기와 6기의 주요 핵심어 중 부산국제영화제, 불꽃축제, MICE산업, 동부산관광단지, 크루즈, 지스타 등은 엔터테인먼트 산업의 메카로서의 부산을 대변해 주고 있으며, 구덕운동장(재건립), 에코델타시티, 산복도로(르네상스)사업 등은 동·서 균형개발의 상징성을 보여주고 있다.

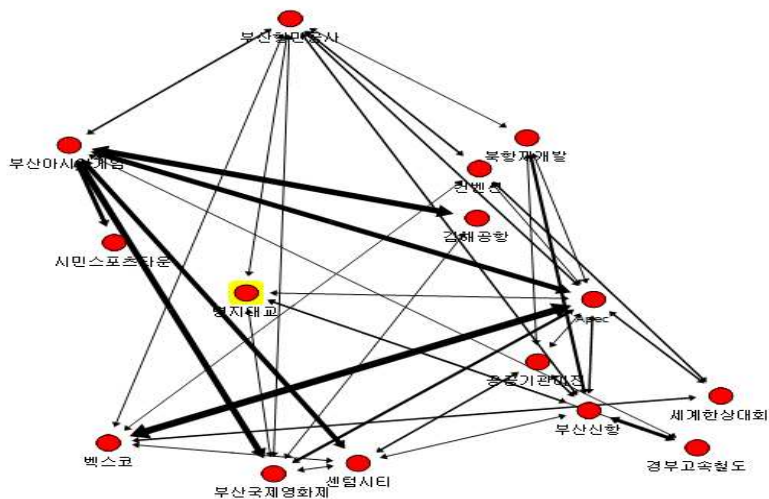
4. 네트워크 분석

지역혁신의 주요 핵심어를 대상으로 민선 3기(2002.7.1. ~ 2006.6.30.) 핵심어간 네트워크 관계를 보면 <그림 2>와 같다. 노드의 갯수는 15개이고 빈도 수가 높을수록 노드가 크다. 노드간의 연계 정도를 나타내는 링크는 굵을수록(weight가 높을수록) 동시출현 빈도가 높다는 의미로 해당 핵심어가 함께 많이 이슈화된 다는 것을 의미한다. <표 2>를 보면 벅스코 및 부산아시아게임이 APEC과 연계가 높게 나타났다. 이는 2002년 부산아시아게임과 2005년 APEC개최가 지역사업에 많이 기여했다는 것을 의미한다.

<표 2> 민선 3기 Weight 높은 순서

node 1	node 2	Weight
APEC	벅스코	19
부산아시아게임	김해공항	16
APEC	부산아시아게임	13
부산아시아게임	부산국제영화제	12
부산아시아게임	센텀시티	8

<그림 2> 민선 3기 네트워크 분석

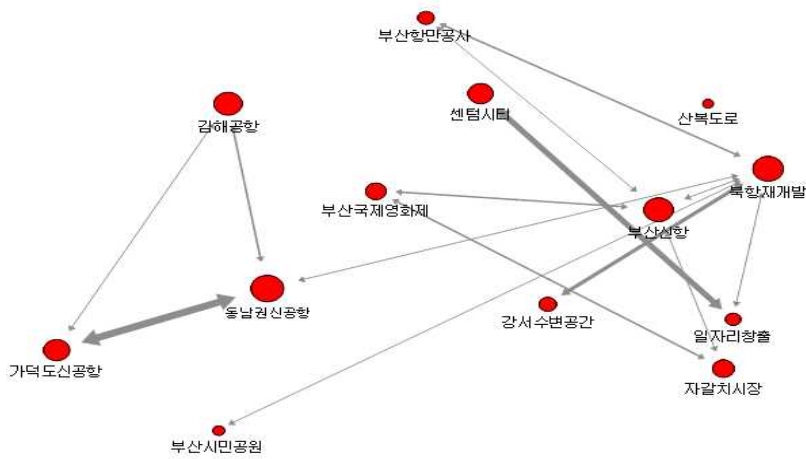


민선 4기(2006.7.1. ~ 2010.6.30.)의 핵심어간 네트워크 관계를 보면 <그림 3>과 같다. <표 3>를 보면 동남권신공항과 가덕신공항건설의 연계가 가장 높게 나타났다. 이는 신공항의 입지 선정문제가 민선 4기 지역혁신의 가장 큰 이슈였다는 것을 보여준다.

<표 3> 민선 4기 Weight 높은 순서

node 1	node 2	Weight
동남권신공항	가덕신공항건설	41
일자리창출	센텀시티	5
북항재개발	강서수변공원	4
동남권신공항	김해공항	3
북항재개발	부산항만공사	3

<그림 3> 민선 4기 네트워크 분석

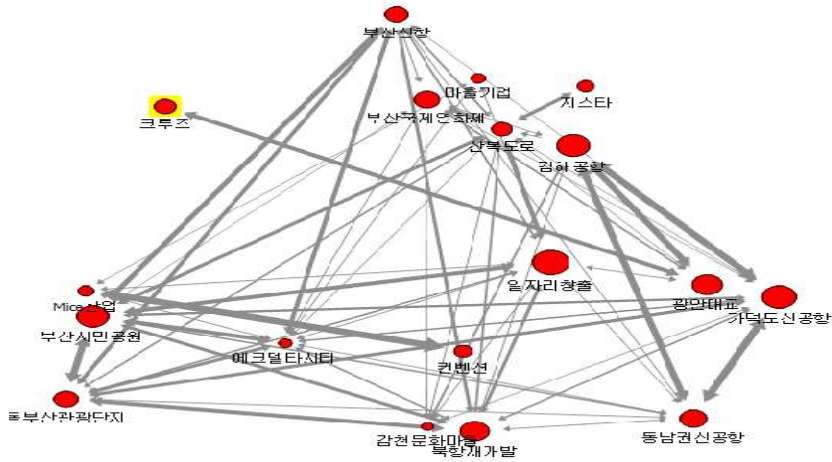


민선 5기(2010.7.1. ~ 2014.6.30.)의 핵심어간 네트워크 관계를 보면 <그림 4>와 같다. <표 4>를 보면 민선 4기와 비슷하게 민선 5기도 신공항의 입지 선정문제가 지역혁신의 큰 이슈였다는 것을 보여준다.

<표 4> 민선 5기 Weight 높은 순서

node 1	node 2	Weight
MICE산업	컨벤션	59
가덕도신공항	김해공항	46
가덕도신공항	동남권신공항	35
김해공항	동남권신공항	23
김해공항	광안대교	18

〈그림 4〉 민선 5기 네트워크 분석

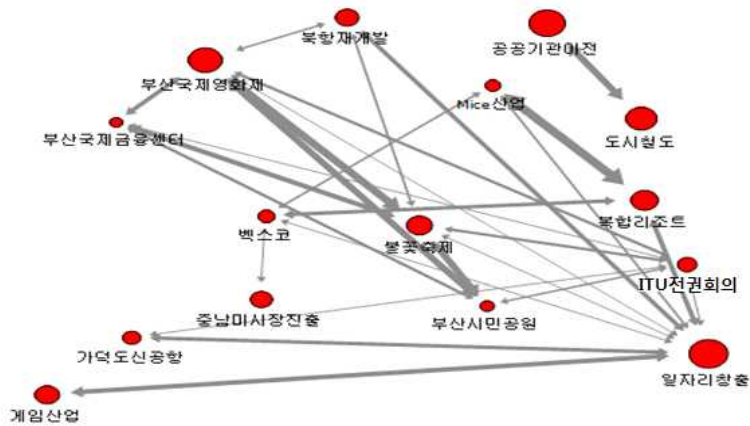


민선 6기(2014.7.1 ~ 2015.10.31.)의 핵심어간 네트워크 관계를 보면 〈그림 5〉와 같다. 〈표 5〉를 보면 MICE산업과 복합리조트건설의 연계가 가장 높고, 다음으로 부산국제영화제와 불꽃축제의 연계성이 높다. 이는 지역사업에 있어서 엔터테인먼트 사업이 점차 커지는 것을 반영하는 것으로 볼 수 있다. 또한 일자리창출과 연계성을 가지는 사업을 보면 북항재개발, 가덕도신공항, 복합리조트사업, 중남미시장진출, 게임산업, MICE산업, 부산국제영화제 등 거의 모든 지역사업과 연관성이 있다. 민선 6기 지역혁신 주요 핵심어에 나타난 사업들은 부산광역시의 5대 도시목표인 문화복지도시, 일자리중심도시, 균형발전도시, 글로벌해양도시, 생활안전도시에 대부분 포함되어 있다. 예를 들면 가덕도신공항건설, 북항재개발, 부산국제금융센터 등은 글로벌해양도시를 만들기 위한 사업들이다.

〈표 5〉 민선 6기 Weight 높은 순서

node 1	node 2	Weight
복합리조트	MICE산업	32
부산국제영화제	불꽃축제	21
공공기관이전	도시철도	15
불꽃축제	부산시민공원	15
부산국제영화제	부산시민공원	14

〈그림 5〉 민선 6기 네트워크 분석



5. 연결 중심성 분석

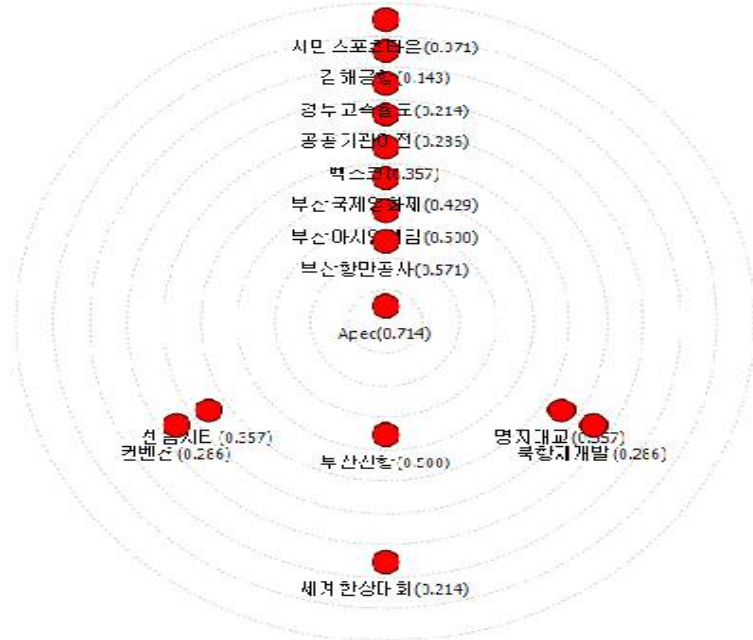
지역혁신 주요 핵심어의 연결정도 중심성(degree centrality)을 민선 3기(2002.7.1. ~ 2006.6.30.), 민선 4기(2006.7.1. ~ 2010.6.30.), 민선 5기(2010.7.1. ~ 2014.6.30.), 민선 6기(2014.7.1. ~ 2015.10.31.)로 구분하여 분석하였다. 연결정도 중심성(degree centrality)은 네트워크의 노드들이 얼마나 많은 연결을 가지고 있는 가를 측정한다. 연결된 노드의 수가 많을수록 연결정도 중심성은 높아지게 된다. 중심성 분석을 통해 한 네트워크에서 중요한 역할을 하거나 주목받는 행위자가 누구인지, 또 각 행위자들은 그 '중심'에 어느 정도 접근하고 있는 지를 알 수 있다(이수상, 2012). 민선 3기에 나타난 주요 핵심어의 연결정도 중심성을 보면 〈그림 6〉와 같다. APEC이 가장 높게 나타났으며, 다음으로 부산항만공사출범, 부산신항건설, 부산아시아게임, 부산국제영화제, 센텀시티조성, 명지대교, 벡스코, 컨벤션, 북항재개발, 공공기관이전 등의 중심성이 높게 나타났다. 대부분 인프라 관련 핵심어의 중심성이 높은 것을 볼 수 있다. 이는 지역혁신을 위해 민선 3기에서는 인프라 구축 관련 사업이 큰 비중을 차지하고 있다는 것을 보여준다. 민선 4기에 나타난 주요 핵심어의 연결정도 중심성을 보면 〈그림 7〉과 같다. 북항재개발이 가장 높게 나타났으며, 다음으로 부산신항, 동남권신공항, 일자리창출, 부산항만공사, 부산국제영화제, 자갈치시장, 가덕도신공항, 김해공항 등이다. 대부분 인프라 관련 핵심어의 중심성이 높은 것을 볼 수 있다.

민선 5기에 나타난 주요 핵심어의 연결정도 중심성을 보면 〈그림 8〉과 같다. 일자리창출이 가장 높게 나타났으며, 다음으로 에코델타시티, 산복도로, 부산신항, 북항재개발, 부산시민공원, 가덕도신공항 등이다.

민선 6기에 나타난 주요 핵심어의 연결정도 중심성을 보면 〈그림 9〉과 같이 일자리창출이 가장 높게 나타났으며, 다음으로 ITU전권회의, 불꽃축제, 부산국제영화제, 부산국제금융센터, 부산시민공원, 벡스코, MICE산업, 북항재개발, 복합리조트건설 등의 연결 중심성이 높게 나타났다. ITU전권회의, 부산국제금융센터, 벡스코, MICE산업 등의 일자리 사업과 부산시민공원, 북항재개발,

복합리조트건설 등의 인프라 관련 핵심어의 중심성이 높은 것을 볼 수 있다. 중심성이 높다는 것은 지역혁신에 있어 민선 6기에서는 일자리 관련 핵심어와 인프라 구축 관련 핵심어가 네트워크에서 큰 비중을 차지하고 있다는 것을 보여준다.

〈그림 6〉 민선 3기 연결 중심성 분석



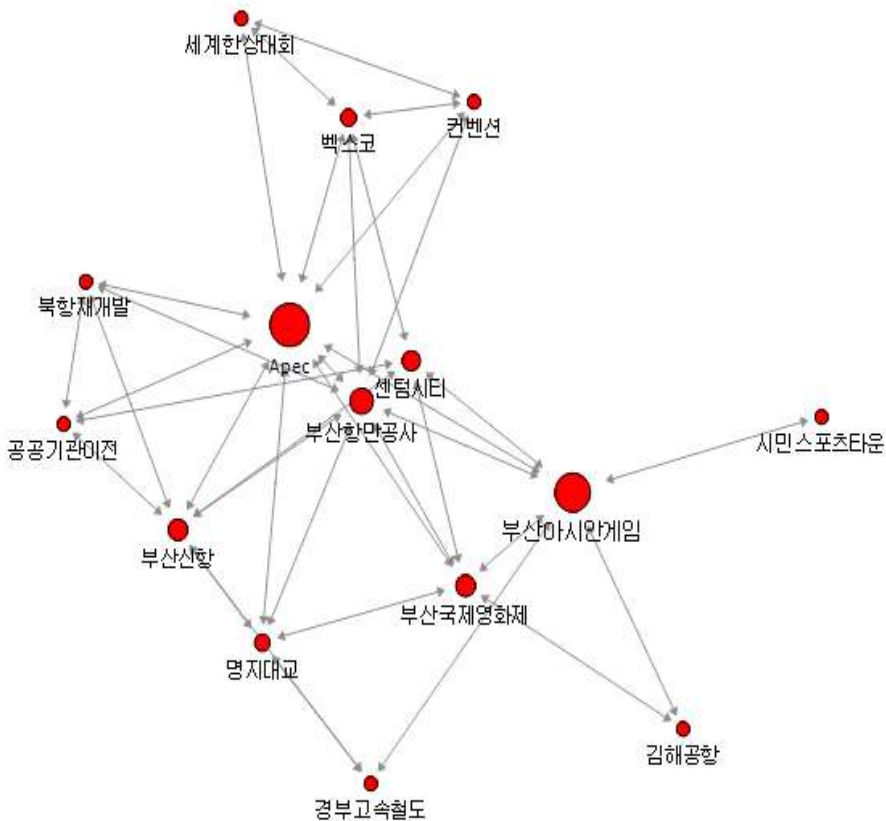
〈그림 7〉 민선 4기 연결 중심성 분석



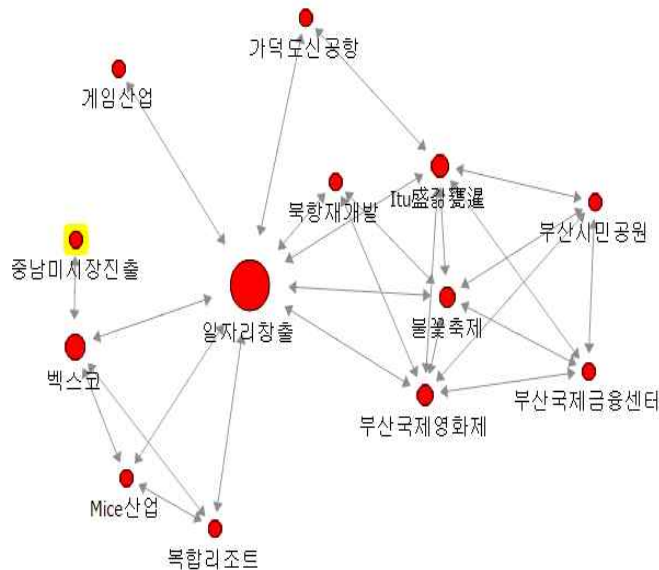
6. 매개 중심성 분석

지역혁신 주요 핵심어의 매개 중심성(betweenness centrality)을 민선 3기(2002.7.1. ~ 2006.6.30.), 민선 4기(2006.7.1. ~ 2010.6.30.), 민선 5기(2010.7.1. ~ 2014.6.30.), 민선 6기(2014.7.1. ~ 2015.10.31.)로 구분하여 분석하였다. 매개 중심성(betweenness centrality)은 한 노드가 다른 노드와 네트워크를 구축하는 데 있어 중개자 혹은 다리 역할을 얼마나 수행하느냐를 측정하는 개념으로 한 노드가 네트워크 내의 다른 노드들의 정보의 흐름을 통제하는 데 큰 영향력을 가질 수 있다(이수상 2012). 매개 중심성이 높은 노드는 중개자의 역할을 한다고 볼 수 있다. 민선 3기, 4기, 5기, 6기에 나타난 주요 핵심어의 매개 중심성을 보면 <그림 10>, <그림 11>, <그림 12>, <그림 13>과 같다. 민선 3기의 매개 중심성은 APEC과 부산아시아안게임이고, 민선 4기의 매개 중심성은 북항재개발, 민선 5기와 6기의 매개 중심성은 일자리창출인 것을 알 수 있다. 매개 중심성이 높은 핵심어는 다른 핵심어들과 연계되는 폭이 넓으며 그 시기의 중요한 화두임을 보여준다. 예를 들어 민선 5기와 6기때의 일자리창출은 다른 사업과 연계되는 폭이 넓으며 중요한 화두임을 의미한다.

<그림 10> 민선 3기 매개 중심성 분석



〈그림 13〉민선 6기 매개 중심성 분석



7. 시기별 지역혁신동향 변화의 특징 및 시사점

본 연구에서는 지역혁신 관련기사를 대상으로 부산시장의 임기에 맞춰 민선 3기에서 민선 6기로 구분하여 분석 하였다. 앞에서 논의한 지역혁신동향 변화의 특징을 종합하면 시기에 따라 지역 혁신 주요 핵심어가 변화하고 있는 것으로 나타났다. 지역혁신 주요 핵심어간 네트워크 관계를 보면 민선 3기의 경우 벅스코와 부산아시안게임이 APEC과 연계가 높게 나타났고, 연결 중심성도 APEC이 가장 높게 나타났으며, 다음으로 부산항만공사출범, 부산신항건설, 부산아시안게임, 부산 국제영화제, 센텀시티조성, 명지대교 등이 높게 나타나 대부분 인프라 사업이 네트워크에서 큰 비중을 차지함을 알 수 있다. 민선 4기의 경우 지역혁신 주요 핵심어간 네트워크 관계를 보면 동남 권신공항과 가덕신공항건설의 연계가 가장 높게 나타났다. 이는 신공항의 위치 선정문제가 민선 4기 지역혁신의 가장 큰 이슈였다는 것을 보여준다. 민선 4기에 나타난 연결 중심성은 복합재개발이 가장 높게 나타났으며, 다음으로 부산신항, 동남권신공항, 일자리창출, 부산항만공사 등이 높게 나타났다. 대부분 인프라 관련 핵심어의 중심성이 높은 것을 볼 수 있다. 민선 5기의 경우 지역 혁신 주요 핵심어간 네트워크 관계를 보면 민선 4기와 비슷하게 신공항의 입지 선정문제가 지역 혁신의 큰 이슈였다는 것을 보여준다. 민선 5기에 나타난 연결 중심성과 매개 중심성은 일자리창출이 가장 높게 나타났다. 민선 6기의 경우 지역혁신 주요 핵심어간 네트워크 관계를 보면 MICE 산업과 복합리조트건설의 연계가 가장 높고, 다음으로 부산국제영화제와 불꽃축제의 연계성이 높다. 이는 지역사업에 있어서 엔터테인먼트 사업이 점차 커지는 것을 반영하는 것으로 볼 수 있다. 민선 6기에 나타난 연결 중심성과 매개 중심성은 일자리창출이 가장 높게 나타났다. 일자리창출과 연관된 핵심어는 여러 분야에서 다양해져 가는 것을 알 수 있다.

본 연구의 시사점은 다음과 같다. 첫째 지역혁신에 있어서 인프라 사업의 중요성이 상대적으로 감소하는 경향을 보이고 있다. 지역혁신 사업을 분석한 결과 인프라 사업은 지역혁신에서 중요한 위치를 차지하고 있으며 특정 시기에 나타나는 사업이라기 보다는 지속적으로 추진되는 사업으로 나타났다. 그러나 최근에는 일자리 사업의 연결 중심성과 매개 중심성이 높게 나타났다. 이는 일자리 창출이라는 핵심어가 다른 핵심어들과 연계되는 폭이 넓으며 중요한 화두임을 보여준다. 지역혁신에 있어서 인프라 사업의 이슈가 상대적으로 감소하고 있는 경향을 보이는 것은 핵심어 분석 결과에도 나타났다. 민선 3기(2002.7.1 ~ 2006.6.30)의 경우 인프라 사업 관련 핵심어의 빈도와 네트워크 연계성, 중심성, 연관성이 높은 것으로 나타났고, 민선 5기(2010.7.1. ~ 2014.6.30.)와 민선 6기(2014.7.1. ~ 2015.10.31.)의 경우 일자리 사업 관련 핵심어의 빈도 및 네트워크 연계성, 중심성, 연관성이 인프라 사업 관련 핵심어와 더불어 높게 나타났다. 이것은 최근 들어 지역혁신에 일자리창출이라는 새로운 이슈가 등장한 것으로, 실업으로 인한 사회적 비용의 급격한 증가는 일자리창출이라는 이슈를 부각시킨 것으로 해석된다.

둘째로 엔터테인먼트 산업 측면에서 지역혁신 사업을 보면 민선 3기 영화산업에서 민선 6기에는 영화산업을 비롯한 MICE산업, 게임산업, 의료관광 등 다양한 사업이 포함된다. 민선 6기의 경우 다양한 사업 관련 핵심어의 빈도 및 네트워크 연계성, 중심성이 높게 나타났다. 부산시는 2004년부터 영상·IT를 4대 핵심 전략산업 중 하나로 선정하여 지속적으로 육성해오고 있다. 특히 부산의 콘텐츠산업 규모는 2013년 기준 업체수는 7,027개, 종사자수는 22,761명이다. 2013년 전국 콘텐츠산업의 총 매출액은 91조912억이며, 이중 서울이 62조1803억으로 68.3%를, 경기도 15조4510억으로 17%를 차지하고 있어 수도권 지역이 콘텐츠산업 매출액 대부분을 차지하고 있다. 부산의 콘텐츠산업 매출액은 2조1129억으로 서울, 경기도 다음으로 높으나 그 비중은 2.3%에 불과하다. 부산의 전국대비 콘텐츠 산업별 매출 비중은 음악산업이 4.0%로 타산업보다 높으며, 이어서 영화(3.8%), 애니메이션(3.3%), 캐릭터(3.2%) 산업 등의 순이다. 전국대비 산업별 매출 비중이 상대적으로 낮은 산업을 살펴보면 만화(1.1%), 지식정보(1.4%), 게임(1.6%), 광고(1.9%), 콘텐츠솔루션(1.9%), 방송(2.0%) 산업 등이다(부산정보산업진흥원, 2015). 부산시는 센텀문화산업진흥지구를 중심으로 부산콘텐츠 제작 기업 및 수도권 역외기업 유치 등 집적화와 함께 콘텐츠산업의 생태계 조성을 위해 노력하고 있다.

셋째로 신공항건설 사업을 보면 민선 3기에서는 신공항 필요성이 일어났고, 민선 4기에서는 신공항 입지 문제가 이슈화되기 시작했으며, 민선 5기에 절정에 이르고, 민선 6기에서는 소강상태를 보여주는 데 이는 지역혁신 주요 핵심어의 빈도 분석과 네트워크 분석 결과로 설명이 된다. 동남권 신공항은 2005년 10월의 영남권 5개 시·도지사의 공동건의에 따라 2006년 12월 정부에서 공식 검토를 시작하였다(김현조 외, 2013). 부산은 가덕도에 24시간 운영 가능한 허브공항 건설을 주장하고 있고 대구·경북·경남·울산 등 4개 지역은 접근성을 내세워 경남 밀양에 신공항 건설을 추진해야 한다고 현재까지 팽팽히 맞서고 있다. 정부에서는 항공수요 조사 결과(“김해공항 항공수요가 2023년 1678만명에 달해 대책이 필요하다”)를 바탕으로 입지 선정을 위한 용역사업의 최종 결과를 2016년 6월에 발표할 예정이다. 동남권 신공항 건설은 부산시민의 염원사업이기 때문에 입지

선정 결과에 따라 부산시장의 리더쉽과 향후 시정에 많은 영향을 미칠 것으로 예상된다.

넷째로 동·서 균형개발 사업은 빈도 민선 5기 이후 활발히 논의되고 있다. 그동안 동부산권의 과다 팽창은 지역 발전에 큰 부담으로 작용했다. 교육·문화·금융·경제 등 모든 분야가 동부산권으로 집중되면서 서부산권 주민들의 박탈감은 심각한 사회 문제로 자리잡았다(신수건, 2013). 이에, 2015년 말 부산시는 서부산의 발전을 부산의 새로운 성장 동력으로 삼겠다는 청사진을 발표하였다. 사상스마트시티, 에코델타시티와 같은 신산업생태 도시공간, 명지를 중심으로 한 R/D거점 지구, 서부산청사와 서부산의료원 건립 등이 대표적 사업이다(서병수 2016). 부산시는 서부산권의 열악한 환경여건의 개선을 통해 산업물류 공간, 주거 공간, 도시 재생, 생태자연환경 등으로 완결되는 글로벌 도시로 도약함으로써 지역발전을 앞당길 수 있을 것으로 기대하고 있다.

다섯째로 동부산관광단지 조성사업을 비롯해 크루즈사업, 복합리조트사업 등 다양한 해양관광 사업이 민선 5기 이후 활발히 논의되고 있다. 세계적으로 전체 관광객중 해양관광객이 차지하는 비중은 계속 증가할 전망이다, 실제 2012년 기준 국내 전체 관광객중 해양관광객은 50% 수준이다(해양수산부, 2014). 부산시 전략경영 100대 추진과제중의 하나로 MICE산업과 해양관광산업의 클러스터 연계방안이 포함되어 있다. 부산지역의 해양레포츠시설과 해양레포츠대회를 MICE산업과 연계해서 지역관광을 활성화하자는 것이다. 해양관광은 특성상 7~8월에 집중되어 있기 때문에 부산시는 나머지 비수기를 활용하기 위하여 부산의 대표적 해수욕장인 해운대 및 광안리, 송도, 다대포해수욕장 등을 문화콘텐츠와 연계하여 사계절 해양관광지로 만든다는 계획이다(해양수산부, 2014).

IV. 결론

지방정부에서는 다양한 사회문제를 해결하기 위해 많은 노력을 하고 있다. 지역문제 해결을 위해서는 문제를 정의하고 범위를 설정해야 한다. 이를 위해 지역혁신정책의 이슈가 어떻게 변화해 왔는 지에 대한 동향분석이 필요하다. 본 연구는 부산시의 지역혁신문제의 동향을 분석하고 다른 지역혁신문제와 연관성이 높고 가장 영향력이 큰 지역혁신문제를 도출하고자 했다. 이를 위해 빅데이터를 활용한 데이터 분석 기법중 비정형 데이터 분석의 하나인 텍스트마이닝 기법을 활용하여 2002년부터 2015년까지 부산 시정신문인 "Bubi News" 기사를 분석하였다. 분석시기는 부산시장의 임기에 맞춰 민선 3기, 민선 4기, 민선 5기, 민선 6기로 구분하였다. 분석 방법으로는 지역혁신 핵심어 빈도분석, 네트워크 분석, 연결 중심성 분석, 매개 중심성 분석을 했다. 본 연구를 통해 지역혁신과 관련된 핵심문제를 도출을 위해 빅데이터 분석 방법을 제시하였고, 후속 연구를 유발하였다는 학문적 기여도가 있다. 또한 실제 지방정부 및 행정기관에서 정책 수립전 지역혁신문제 중 이슈가 되는 사업에 대한 전체 흐름을 파악하기 위한 자료로 활용할 수 있다는 실무적인 기여도가 있다. 하지만 본 연구에서는 부산시 지역혁신문제를 도출하기 위해 시정신문인 "Bubi News" 기사만을 선정하여 분석하였기 때문에 전반적인 전체 동향을 반영함에 있어 제한점이 있을 수 있다. 향후 이러한 점을 보완하기 위해 "Bubi News" 외에도 부산 지역 신문기사들을 추가검색 및 분

석할 필요가 있다. 또한 주제와 관련된 뉴스 기사를 추출하는 과정에서 주관성을 감소시켜야 할 필요성이 있다. 어떤 연관키워드를 대상 주제로 선정하는가에 따라 추출되는 뉴스 기사의 방향성은 크게 달라질 수 있다. 본 연구에서는 관련문헌 검토와 혁신 전문가 2명과 정책 전문가 1명이 브레인스토밍을 통하여 연관키워드 선정과 빈도 분석전에 명사화하는 1차 작업 및 선정된 단어들의 상대비교 수준을 맞추기 위한 2차 작업을 하였으나, 추후 좀 더 구체적이며 객관적인 방법론을 적용한다면 더 타당성있는 결과를 도출할 수 있을 것이다.

참고문헌

- 김주식·김무용·현병환. (2013). 특허정보 기반의 바이오 기술개발 트렌드 분석 및 유망기술분야 도출에 관한 연구. 「기술혁신연구」, 21(2): 25-56.
- 강범모. (2012). 가치있는 정보를 찾아주는 텍스트마이닝과 신문. 「동아일보」,
<http://news.donga.com/3/all/20120410/45407849/1>
- 구주나. (2014). 「한국어 텍스트 마이닝의 특징 고찰 및 실제 빅데이터에의 적용」. 석사학위논문. 숙명여대
- 김정홍. (2005). 「기술혁신의 경제학」. 시그마프레스
- 김하진·송민. (2014). 동시출현단어 분석을 통한 국내외 정보학 학회지 연구동향 파악. 「정보관리학회지」, 31(1): 99-118.
- 김현조·허철행·이수구·박영강. (2013). 동남권 신공항의 정책목표와 새로운 접근방법. 「지방정부연구」, 16(4): 7-32.
- 두산백과. <http://www.doopedia.co.kr/search/encyber/totalSearch.jsp?WT.ac>
- 박두순·문양세·박영호·윤찬현·정영식·장형석. (2014). 「빅데이터 컴퓨팅 기술」, 아카데미
- 박한우·Loet Leydesdorff. (2004). 한국어의 내용분석을 위한 KrKwic 프로그램의 이해와 적용: Daum.net에서 제공된 지역혁신에 관한 뉴스를 대상으로. 「Journal of The Korean Data Analysis Society」, 6(5): 1377-1388.
- 부산정보산업진흥원. (2015). 「2014 부산CT산업 현황조사」, 17.
- 신수건. (2013). 서부산권 지도 바뀌놓을 메가 프로젝트 속속 진행. 「부산발전포럼」, 3(4): 1-8.
- 서병수. (2016). 서병수 부산시장 2016년 신년사.
<http://blog.naver.com/suhbs8680/220586938414>
- 오상영. (2015). 위기의 지방과 기업투자 유치. 「중부매일」,
<http://www.jbnews.com/news/articleView.html?idxno=686154>
- 이승현. (2012). 「기술혁신기업의 마케팅혁신에 관한 연구 : 대기업과 중소기업간의 비교분석」. 석사학위논문. 건국대학교
- 이수상. (2012). 「네트워크 분석방법론」, 39-41, 255-267. 논형
- 이인석·박승환·백준걸. (2015). 텍스트마이닝을 통한 제조시스템 분야의 연구동향파악. 「대한산

- 업공학회 춘계공동학술대회 논문집, 4201-4205.
- 양현수. (2012). 「텍스트 마이닝을 활용한 청소년 범죄 및 비행의 유형 변화 연구」. 석사학위논문. 경상대학교
- 임시영·이미숙·진기호·신동빈. (2014). 텍스트마이닝 기술을 이용한 공간정보 분야의 연구 동향에 관한 고찰 - 국가연구개발 사업 보고서 및 논문을 중심으로. 「한국공간정보학회지」, 22(4): 11-20.
- 임성배·홍순구·김주남. (2014). 창조경제 성공을 위한 프레임과 공동혁신방안에 관한 연구. 「방통융합기반정책연구 14-03」. 미래창조과학부.
- 원진영·김대곤. (2014). 텍스트마이닝을 활용한 사회위험 이슈 도출. 「한국위기관리논집」, 10(7): 33-52.
- 조운애·김원규·남장근·오준병. (2005). 「혁신역량 강화를 위한 연구개발투자의 효율성 제고방안」. 산업연구원
- 정다운·유지송·이미숙·신동빈. (2015). 텍스트 마이닝에 기반한 U-City 서비스 이슈 및 동향분석 - 네트워크 분석 및 정보량 계측기법을 활용하여. 「한국공간정보학회지」, 23(5): 35-44.
- 해양수산부. (2014). 「제2차 해양관광진흥기본계획」, 8-10, 26-27.
- 최영출·박수정. (2011). 한국행정학의 연구경향분석: 네트워크 텍스트 분석 방법의 적용. 「한국행정학보」, 45(1): 123-139.
- Feldman, R. & Dagan, I. (1995, August). Knowledge Discovery in Textual Databases (KDT). In *KDD 95*: 112-117.
- Hotho, A. Nürnberger, A., & Paaß, G. (2005, May). A Brief Survey of Text Mining. In *Ldv Forum 20*(1): 19-62.
- Mohamed M. Mostafa. (2013). More than Words: Social Networks' Text Mining for Consumer Brand Sentiments. *Expert Systems with Applications*, 40(10): 4241-4251.
- Schumpeter, J. A. (1934). *The Theory of Economic Development*, Harvard University Press. ; Myers, S. and D. G. Marquis (1969), *Successful Industrial Innovations: A Study of Factors Underlying Innovation in Selected Firms*, *National Science Foundation*, 재인용

박주섭(朴柱燮): 동아대학교 경영정보학 박사과정에 재학중이고 학문적 관심분야는 Co-creation, 지역혁신이다(juseop60@naver.com).

홍순구(洪淳玖): University of Nebraska-Lincoln에서 경영정보학 박사학위를 취득하고 동아대학교 경영정보학과 교수로 재직중이다. 학문적 관심분야는 Co-creation, 웹 접근성, 정보시스템 평가, 지역혁신 등이고 주요저술과 발표논문으로 「유비쿼터스의 이해와 활용사례, 정보기술의 전략적 활용: 경영정보시스템」, 「Co-creation 플랫폼 운영 전략: 삼성 앱스와 애플 앱스토어 비교(2014)」, 「ICT기반 지역공동체 가치창출 전략의 탐색」(2013), 「지방자치단체의 모바일 앱 서비스 이용 활성화 방안에 관한 연구: 부산광역시를 중심으로」(2013) 등이 있다. 주요경력으로는 Texas A&M International University 조교수, 한국은행 근무, 동아대학교 재무처장 등이다(shong@dau.ac.kr).

Abstract

Research of Trends in Busan Regional Innovation Policies Using Text Mining

Park, Ju Seop
Hong, Soon Goo

In the modern society, various social problems are occurring and regional governments are stepping up their efforts to solve these problems. To find out the solution, it is preferentially necessary to grasp the present social problems that a region is facing. Especially, it is important to examine the trend of policies for the regional innovation related to the local economic growth and development. The purpose of this study is to analyze the trend of policies for the regional innovation and draw regional innovation policies highly related with the others and more influential than them. To this end, we utilized text mining method, which is one of the informal data analysis method and analyzed the articles from 2002 to 2015 of "Buvi News"- Busan municipal government newspaper. The results of analysis showed the following implications. The importance of the infrastructure issue for the regional innovation is relatively decreased, while entertainment business has been a hot issue. In addition, it has been actively discussed on the activation of marine tourism and balanced development of the east and west area in Busan since the 5th mayor's term elected by popular vote. This study suggested a analysis method to draw a major issue related with the regional innovation and gave an academic contribution, triggering a follow-up study. Practically, the result of this study could be utilized as guideline for the policy establishment in the local government and the administrative agency.

Key Words: Regional innovation policies, Network analysis, Text mining, Structural analysis, Tendency analysis