



ISSN 1229-8565 (print)

한국지역사회생활과학회지

Korean J Community Living Sci

<http://dx.doi.org/10.7856/kjcls.2016.27.S.665>

ISSN 2287-5190 (on-line)

27(S): 665~679, 2016

27(S): 665~679, 2016

## 빅데이터 분석을 통한 농촌관광 실태와 활성화 방안 연구: 전라북도를 중심으로

박 로 운<sup>1)</sup> · 이 기 훈<sup>2)†</sup>

전주대학교 대학원 생명자원융합과학과<sup>1)</sup> · 전주대학교 경영학과<sup>2)</sup>

### Study of the Activation Plan for Rural Tourism of the Jeollabuk-do Using Big Data Analysis

Ro Un Park<sup>1)</sup> · Ki Hoon Lee<sup>1),2)†</sup>

Dept. of Integrated Bio-Resource Science, General Graduate School of Jeonju University, Jeonju, Korea<sup>1)</sup>

Dept. of Business Administration, Jeonju University, Jeonju, Korea<sup>2)</sup>

#### ABSTRACT

This study examined the main factors for activating rural tourism of Jeollabuk-do using big data analysis. The tourism big data was gathered from public open data sources and social network services (SNS), and the analysis tools, 'Opinion Mining', 'Text Mining', and 'Social Network Analysis(SNA)' were used. The opinion mining and text mining analysis identified the key local contents of the 14 areas of Jeollabuk-do and the evaluations of customers on rural tourism. Social network analysis detected the relationships between their contents and determined the importance of the contents. The results of this research showed that each location in Jeollabuk-do had their specific contents attracting visitors and the number of contents affected the scale of tourists. In addition, the number of visitors might be large when their tourism contents were strongly correlated with the other contents. Hence, strong connections among their contents are a point to activate rural tourism. Social network analysis divided the contents into several clusters and derived the eigenvector centralities of the content nodes implying the importance of them in the network. Tourism was active when the nodes at high value of the eigenvector centrality were distributed evenly in every cluster; however the results were contrary when the nodes were located in a few clusters. This study suggests an action plan to extend rural tourism that develop valuable contents and connect the content clusters properly.

**Key words:** rural tourism, big data, Jeollabuk-do

Received: 4 October, 2016 Revised: 19 October, 2016 Accepted: 24 October, 2016

†Corresponding Author: Ki Hoon Lee Tel: +82-63-220-2279 E-mail: khlee@jj.ac.kr

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

## I. 서론

2016년의 우리나라의 농촌 관련 국정과제를 살펴보면 농촌관광 소득화 기술개발, 농촌관광 상품개발, 농업유산 관리 및 관광자원 기술개발 등으로 농촌의 6차 산업화를 촉진시키기 위한 지원이 활발하게 진행되고 있다. 그 중에 농촌관광개발은 관광시장을 파악하기 위한 인프라를 구축하고 효율적인 마케팅 전략을 수립한 지역 또는 중앙정부가 농촌 간의 관광개발 경쟁에서 우위를 차지할 수 있기 때문에(Kim 2015) 각 지역은 관광시장에 대한 다양한 분석과 함께 관광활성화 전략 수립에 힘쓰고 있다.

우리나라 농촌지역에는 정부주도하의 농촌관광개발이 시작된 이래 2002년부터 중앙정부 및 시군지자체를 통틀어 1,000개 이상의 농촌관광마을이 개발 중이다(Yoon et al. 2009). 농촌지역은 자연환경을 기반으로 하여 문화역사, 엔터테인먼트, 음식, 축제, 체험 등과 같은 콘텐츠를 통해 방문객을 유치하기 위한 지역 개발을 하고 있다(Park 2015).

그러나 최근 우리나라의 농촌지역은 국제정세 및 장기적 경제난에 따른 지역산업의 불황, 농산물의 수입개방으로 인한 경쟁력 저하, 농촌인구의 고령화 등에 의한 지역의 성장저하가 큰 문제로 야기되고 있다(Park 2015). 일부 선진국들은 농촌관광을 통한 지역의 활성화가 이루어지고 있는 것으로 나타났으며 우리나라는 농촌관광개발사업을 추진하고 있지만 경영능력 부족, 관광지역의 낮은 경쟁력, 과도한 시설투자 등으로 성과가 미비한 실정이다(Park 2015). 또한 유럽 선진국의 경우 농촌관광이 국내 관광의 약 20%를 차지하고 있는 데 비해, 우리나라의 농촌관광은 국내 관광의 약 3.2%를 차지하고 있을 뿐이다(Park 2013). 우리나라의 농촌관광은 농촌관광의 공급자 수가 증가하면서 농촌이 활성화되어가고 있지만, 관광 콘텐츠의 품질 저하 및 중복, 운영 전문성 부족, 지역주민과의 갈등으로 인한 문제점도 나타나고 있는 추세이다(Agricultural Research & Extension Services 2010). 따라서 이러한 농촌의 문제점을 해결하고 농촌지역의 관광을 활성화

하는 방안에 관한 연구의 필요성이 있다.

전라북도는 시·군별 자연자원과 다양한 문화역사자원을 보유하고 있고 식문화가 발달한 편이다. 관광패턴이 다양화되면서 농촌관광의 수요가 증가하고 익산과 장수를 잇는 고속도로 등으로 접근성 상승과 새만금사업, 군산공항 국제선취항, 익산 푸드클러스트 등의 콘텐츠요소가 방문객을 유치하는 새로운 요인으로 자리잡고 있다(Agricultural Research & Extension Services 2010). 체험에 관련한 방문객이 많고 농촌관광에 어울리는 친환경 농업과 생명산업이 발전하였지만 농촌관광 경영체 간의 공동체의식이 결여되어있고 지역 내 농촌관광 콘텐츠 간의 연결성은 미흡한 수준이다(Agricultural Research & Extension Services 2010). 14개의 시·군으로 구성되어 있는 전라북도지역의 주요관광지 방문객은 2015년을 기준으로 약 2,900만 명으로 지역별 분포도 편차가 없는 편인데 이는 각 지역별 우수한 농촌관광 경영체들이 산발적으로 존재하고 있기 때문이다(Agricultural Research & Extension Services 2010). 전라북도는 농촌관광의 활성화를 통해 지역경제를 살리고자하는 의지를 가지고 있지만 각 지자체별로 난립하는 축제와 관광자원 개발 등만으로는 그 성과를 기대하기 어려운 실정이다. 그런데 최근에 관광객들에 대한 정보가 축적되면서 이를 활용하여 우수한 관광프로그램의 개발이나 관광객 유치증대를 이룰 수 있는 환경이 조성되었다. 특히 SNS(Social Network Services)를 통한 검색이나 키워드 등에 의하여 관광객유인 요인이나 불만족사항 등의 문제점들을 파악하기 용이하게 되었고 이러한 빅데이터 분석을 통하여 농촌관광의 활성화 전략수립이 가능하게 되었다.

국내의 농촌관광산업 분야에서 빅데이터를 활용하고 관련 통계를 활용한 연구는 초기 단계에 머물러 있으며 특히 관광수요를 예측하고 관광지 활성화 방안에 관한 통계를 이용한 분석 역시 미비한 실정이다(Kim 2015). 관광이 기존의 산업범위의 한계를 탈피하여 타 산업과의 융합을 통한 새로운 비즈니스 모델이 창출되어야 하는데(Kim et al. 2014) 관광기업들이 아직 빅데이터를 운영할만한 하드웨어 및 소프트웨어를 갖추지 못한 편이고 타 분야와의 연결성도 낮은 편이다. 따라

서 관광산업 분야에서 관련분야 빅데이터를 활용한 농촌지역 관광지 활성화에 대한 방안에 관한 연구의 필요성이 제기된다.

따라서 본 연구에서는 여러 기관에서 축적한 빅데이터를 이용하여 전라북도 14개 지역의 농촌관광에 대한 현황과 문제점 등에 관한 분석을 실시하였다. 이를 통하여 새로운 관광유인 요인이 조성되고 각 관광 경영체들이 존재하고 있지만 그 연결성이 부족하여 농촌관광이 활성화되고 있지 않다는 기존연구의 판단에 대해서도 데이터에 의한 검증을 실시할 것이다. 또한 데이터분석결과를 바탕으로 농촌관광의 활성화 요인을 제시함으로써 전라북도 지역의 경제적 자립과 발전에 기여하고자 한다.

## II. 이론적 배경

### 1. 관광산업 분야 빅데이터 연구 동향

빅데이터(big data)는 거대한 규모의 자료를 의미하는 용어로 대용량의 데이터를 분석, 추출, 예측, 활용하기 위한 기술이다(Oh et al. 2015). 빅데이터는 IT기술의 발전, 소셜미디어의 발달 등으로 대규모 데이터들이 발생하고 저장되면서 기존의 데이터베이스로는 처리하기 어려워진 문제를 해결하기 위해 여러가지 추출기술이 개발되면서 시작되었다. 또한 추출된 빅데이터에 대한 데이터 처리기술이 발달하고 처리속도가 빨라짐에 따라 다양한 분야에서 활발하게 이용되게 되었다. 최근에는 정부의 공공데이터 개방과 함께 관광산업 분야에서의 빅데이터가 활용되고 있는데 관광산업 분야에서는 스마트기기의 증가와 정보통신기기가 발달하면서 생성되는 실시간 관광데이터를 바탕으로 방문객의 동기 및 행태 등을 분석할 수 있게 되었다(Kim 2015). 국내외 관광산업 분야에서 빅데이터와 관련된 연구동향은 다음과 같으며 이는 Table 1에 정리되어 있다.

#### 1) 국내 관광산업 분야 빅데이터 연구동향

Kong & Song(2013)은 호텔의 블로그 등의 SNS상

의 빅데이터에서 해당 호텔관련 키워드에 관한 고객의 의견들을 수집하고, 분석하여 호텔 CRM(Customer Relationship Management)의 개선점을 찾아내고 개선하는 방법을 제안하였다. 서비스 모델을 정립하여 미디어채널, 소비자만족도, 브랜드이미지 등을 분석하여 여행사가 효율적으로 고객을 관리하는 방법을 제안하였다. Lee & Yoon(2014)은 2013년 SNS를 실시간으로 분석해주는 다음소프트의 'Social Metrics'를 이용하여 특정 관광 검색 키워드와 관련된 연관 키워드를 분석하여 의미를 도출하는 탐색적 연구를 진행하였다. 여기서는 빅데이터를 관광연구에서 어떻게 활용할 것인지에 대한 지침을 제공하였고 비정형의 자료를 의미 있는 자료로 해석하고 현장 적용시키는 것에 대한 기본틀을 제시하였다. 여기서는 연관성이 높은 키워드를 선정하여 관광지와 숙박시설에 대한 마케팅 시사점을 제안하였다. Jung(2015)은 SNS 고객성향 분석을 통해 사용자의 감성적 측면을 포함한 여행지 선별법을 제안하였다. 텍스트 데이터에 내재되어진 시멘틱(semantic) 패턴을 찾아내는 기법을 활용하여 고객성향을 바탕으로 최적의 여행지 조건을 제시하고자 하였다. 텍스트 뿐 아니라 비정형 데이터인 이미지, 멀티미디어, 소셜미디어 등의 분석에서도 널리 쓰이고 있는 토픽모델링 중 하나인 LDA(Latent Dirichlet Allocation) 방법과 MCMC 기법을 이용하여 토픽과 감성(sentiment)을 같이 다루는 기법을 제안하였다. Jeon(2015)은 소셜미디어에 나타난 한국철도공사의 상품인 '내일로'의 이용후기 1050건을 수집하여 연관성규칙 분석과 키워드 네트워크 분석을 시행하였다. 연관성 분석을 통하여 소비자들의 이용패턴을 파악할 수 있는 역들 간의 104개의 법칙을 도출하였다. 소비자들의 행동 및 패턴분석, 수요예측을 통하여 기업의 경영성과 향상 및 서비스 개선 경쟁력 강화가 가능함을 보여주었다. Cho et al.(2015)은 웹크롤러를 이용하여 뉴스, 블로그 등으로부터 충청북도 관광데이터를 수집하고 공공데이터와 통합하여 텍스트마이닝 기법으로 분석하여 충북관광을 위한 홍보 전략과 관광진흥정책을 수립하는 방안을 제시하였다. 분석결과 충청북도 북부지역이 주요 이슈 관광지로 나

**Table 1.** Summary of big data research for tourism

Researcher	Data source	Methodology	Results and applications
Lau et al. 2005	Blog	Text mining	<ul style="list-style-type: none"> <li>- Show images of tourists for hotels of Hong Kong</li> <li>- Suggest a practical marketing plan</li> </ul>
Jeong & Jeon 2008	TripAdvisor review	Text mining ANOVA	<ul style="list-style-type: none"> <li>- Draw satisfaction and intention to revisit as major reward factors</li> <li>- A daily rate affects mainly satisfaction and revisit</li> </ul>
Ye et al. 2009	Yahoo travel site	Discriminant analysis	<ul style="list-style-type: none"> <li>- Construct modeling for positive/negative emotion for tour place</li> </ul>
O'Connor 2010	TripAdvisor review	Content analysis, frequency	<ul style="list-style-type: none"> <li>- Investigate satisfactions with reviews</li> <li>- Managing reviews matters</li> </ul>
Kong & Kang 2013	Blog, Facebook, Twitter	Keyword, text mining	<ul style="list-style-type: none"> <li>- Apply hotel CRM &amp; marketing</li> <li>- Offer data security improvement</li> </ul>
Duan et al. 2013	TripAdvisor review	Correlation, regression	<ul style="list-style-type: none"> <li>- Amounts of reviews are positive</li> <li>- Positive reviews affects resales</li> </ul>
Lee & Yoon 2014	Facebook, Twitter	Keyword, social metrics	<ul style="list-style-type: none"> <li>- Analyse unstructured data</li> <li>- Distract meanings from tourism keywords</li> </ul>
Jung 2015	Facebook, Twitter	Latent Dirichlet Allocation(LDA), MCMC	<ul style="list-style-type: none"> <li>- Recommend consumers customized tour place</li> <li>- Find semantic pattern of data</li> </ul>
Xiang et al. 2015	Expedia review	Text mining	<ul style="list-style-type: none"> <li>- Classify keywords with satisfaction</li> <li>- Confirm the relation of experience and satisfaction with consumer act</li> </ul>
Cho et al. 2015	News, Blog, Webpage	Web Crawler, text mining	<ul style="list-style-type: none"> <li>- Perform emotional analysis for tour data of Chungbuk with keywords</li> <li>- Establish tour strategies and policies</li> </ul>
Kim 2016	TripAdvisor review	Text mining, SPSS, R	<ul style="list-style-type: none"> <li>- Find factors for consumer's choice</li> <li>- Suggest marketing plan for small hotels</li> </ul>
Sim 2016	Korea Tourism Org. review	Regression, Factor analysis	<ul style="list-style-type: none"> <li>- Find factors for image of tour place and relations with satisfaction</li> <li>- Produce resources for regional marketing strategy</li> </ul>

타났으며, 주요 관광지별 감성분석을 통하여 긍정 및 부정 현황과 원인을 파악하였다. 또한 마케팅전략을 수립하는데 활용할 수 있도록 다차원분석기법을 이용하여 관광관련 텍스트 정보를 다각도로 분석하였다. Kim(2016)은 TripAdvisor의 평판 중 서울의 호텔 총 131개 중 82개 표본을 추출하여 텍스트마이닝을 실시하여 이용자중심의 호텔선택속성을 도출하는 연구를 진행하였다. 중저가호텔의 만족도를 성별, 나이 등을

기준으로 조사하여 중저가호텔의 경쟁력을 갖출 수 있도록 마케팅방향을 제시하였다. Sim(2016)은 한국관광공사의 외국인 리뷰데이터를 수집하여 관광지 이미지 구성요인과 만족도 간의 회귀분석을 통하여 상관관계를 파악함으로써 관광지 이미지 구성요인 도출과 관광지 이미지 결정 요인을 연구하였다. 연구결과, 관광지의 8개의 인지적 이미지요인과 8개의 정서적 이미지요인이 도출되었으며, 도출된 관광지 이미지는 관광지

유형별로 유의한 차이가 있는 것으로 나타났다. 또한 관광지 이미지와 관광지 유형의 연관성 요인을 밝혔으며 관광지 이미지와 만족도 간의 유의한 상관관계를 확증하였다.

이와 같은 관광관련 키워드를 이용한 속성이나 이미지 분석 외에 시스템에 관련한 연구도 존재하였다. Lee et al.(2010)은 N포털 사이트를 대상으로 포털검색 엔진의 공공적인 역할 수행에 대해 분석하였다. Kim et al.(2014)은 OPEN-API를 이용하여 제주관광 관련 키워드 추출 시스템을 개발하였고 소셜데이터 기반의 제주관광 트렌드 서비스를 개발하였다. Yoo(2014)는 비정형화 SNS 사이트를 통하여 병렬형 검색과정을 수행한 뒤 집중형 검색과정으로 강한 신뢰형 감성정보를 기반으로 DB화 및 시각화 서비스로 연계자료를 제공하였다(Kim 2015; Kim 2016).

## 2) 국외 관광산업 분야 빅데이터 연구동향

국외 연구도 후기나 SNS 자료를 통한 텍스트마이닝 분석과 기초적인 통계분석에 한정되어 있다. Lau et al.(2005)은 블로그에서 수집한 데이터를 바탕으로 여행객들이 홍콩호텔에 대한 인식을 분석하고 이를 홍콩호텔의 마케팅에 활용할 수 있는 방안을 제시하였다. Ye et al.(2009)은 야후(yahoo.com)의 관광사이트인 travel.yahoo.com에서 관광리뷰 데이터를 긍정과 부정의 감정형태로 구분하는 기법을 제안하고 이에 관한 모델링 방안을 제시하였다. Duan et al.(2013)은 TripAdvisor의 호텔후기를 감정평가기법을 통해 분석하여 긍정적인 후기가 방문자의 기대수준을 높여 궁극적으로 방문자수의 증가에 영향을 주므로 성공적인 호텔경영을 위해서 후기관리가 필요하다고 제안하였다. Xiang et al.(2015)은 Expedia.com의 소비자 리뷰를 바탕으로 호텔후기 및 평가가 소비자행동에 영향을 주고 있음을 밝혔다. 고객의 평가를 기반으로 호텔선택속성과 만족도를 도출하였다. Li et al.(2015)은 홍콩의 39개의 호텔의 후기를 통해 호텔선택 및 평가의 요소에 대한 연구를 하여 호텔의 시설과 주변 환경, 교통 등이 호텔선택에 영향을 주는 요인으로 제시하였다. 이 연구는

텍스트마이닝 기법을 사용하였고 결과로는 호텔의 기본적인 호텔 시설뿐만 아니라, 주변의 환경과 교통기능도 고객의 평가에 중요한 영향을 미침을 규명하였다. Jeong & Jeon(2008)은 TripAdvisor를 통해 뉴욕 소재의 호텔 324곳의 후기를 추출하여 방문객의 만족평가와 재방문의사에 중요한 9가지 속성을 찾아내고 전체적 만족도와 재방문에서 일일평균가격이 가장 큰 영향을 주고 있음을 밝혔다. O'Connor(2010)도 영국의 100개 호텔을 선정하여 TripAdvisor를 통해 관련 평가데이터를 추출하여 분석함으로써 평가데이터를 활발히 관리하는 것이 매우 중요하다는 사실을 밝혔다. Lee et al.(2011)은 Daodao에서 베이징의 1063곳의 호텔에서 리뷰데이터를 수집하여 내용분석을 하여 호텔의 가격 수준에 따라 고객 만족의 결정요인 15가지를 밝혔다. Schuckert et al.(2015)은 TripAdvisor의 평가 데이터를 통해 홍콩의 185개의 호텔의 평점과 개인평가의 관계를 밝혔는데, 등급이 낮은 호텔에서 평점의 편차가 커서 개인들 간의 평가점수가 일정하지 않음을 보였다(Kim 2015; Kim 2016; Sim 2016)

## 2. 기존 연구의 한계

현재 우리나라의 많은 산업분야에서 빅데이터를 활용하여 미래를 예측하고 가치를 창출하는 유용한 사실을 발견하고자 하는 빅데이터 연구들이 활발히 진행되고 있다. 그러나 관광분야에서의 국내의 연구는 빅데이터 관련 인프라가 잘 구축되어 있지 않은 실정이라 초기단계에 머물러 있다고 할 수 있다.

기존연구의 문제점을 살펴보면 첫째, 국내의 관광분야에 활용된 빅데이터의 경우 관광을 대상으로 한 연구들과 소셜미디어를 활용한 관광 빅데이터 분석의 수준이 아직 미비한 수준이기 때문에, 관련연구들이 추가로 이뤄져야 할 필요성이 제기되고 있다(Kim 2015). 빅데이터를 활용한 관광 분야의 기존 연구들은 대부분 SNS를 이용한 데이터분석 또는 여행지관련 후기(review) 등을 통해 텍스트마이닝을 하였다. 즉, 분석방법과 분석자료에 대한 다양성이 부족하다고 할 수

있다. 둘째, 관광 분야에 활용한 빅데이터에 관련한 기존의 연구들은 표본데이터의 수가 비교적 적다. 빅데이터 분석이라 하기에는 표본수가 많지 않아 그 결과의 신뢰성을 담보하기에 어려움이 있는데 이는 빅데이터 수집을 권장하는 기관의 부족, 자료공유의 제한 등에 기인한다고 할 수 있다. 요약하면 관광분야를 연구하기 위해 정부에서 제공하는 공공데이터가 한정적이고 일반기업체의 관련자료 취득도 용이하지 않기 때문에 보다 폭넓은 연구를 수행하는데 한계가 있다는 것이다. 이러한 문제점은 국내연구에서만 나타나는 현상이 아니고 해외연구에서도 발생하는데 주로 후기를 대상으로 하는 텍스트마이닝 분석방법에 주력하고 있어 연구결과도 후기가 좋으면 만족도가 높다는 등의 단순한 결론에 그치는 경우가 대부분이다.

따라서 본 연구에서는 빅데이터를 활용한 관광분야에서 SNS를 통한 관광지 또는 편의시설, 여행사 등의 평가의 데이터분석 및 여행지관련 리뷰 등을 분석한 기존의 연구들과는 달리 국가가 제공하는 공공데이터 중 농촌지역관광지 방문객데이터와 콘텐츠분포에 따른 농촌지역 축제방문객 수요를 통해 기초적인 통계분석을 실시한 후, 오피니언 마이닝을 통한 'Social Metrics'와 텍스트마이닝을 통한 'Textom'을 이용해 콘텐츠분석을 진행하고 소셜네트워크분석(Social Network Analysis, SNA)의 시각화도구인 'Gephi'를 통해 각 지역의 콘텐츠의 분포 및 연결성을 분석하였다. 이러한 연결성을 근거로 하여 문제점과 시사점을 도출하고, 농촌관광 활성화를 위한 방안을 모색하여 결론으로 제안하였다.

### III. 연구방법

#### 1. 분석도구

본 연구에서는 SNS상의 관광관련 데이터를 사용하여 오피니언마이닝(Opinion Mining)과 텍스트마이닝(Text Mining), 그리고 소셜네트워크분석(SNA)을 진행하였다.

오피니언마이닝을 위한 빅데이터 분석도구는 다음 소프트웨어(www.daumsoft.com)에서 제공하는 'Social Metrics'

로 트위터와 블로그 자료의 연관 키워드, 감성 키워드, 주간 급증 키워드 순위 및 추이를 검색할 수 있고 탐색어 여론 및 긍부정에 관한 추이를 제공한다. 본 연구에서는 Social Metrics의 탐색어 맵의 결과를 바탕으로 분석하였고 검색어는 전라북도 14개 지역을 사용, 'OO 여행'으로 하였고 연관어 수는 15개로 지정하였다.

텍스트마이닝을 위한 빅데이터 분석도구는 'Textom'으로 웹(web)과 SNS상의 데이터를 검색할 수 있고 수집데이터 및 보유데이터의 처리도 가능한 2-way 정제 분석 기능을 제공한다. 또한 차트 및 그래프를 직관적이고 심미적으로 표현하며 맞춤형 데이터 정제 및 다양한 분석프로그램(UCINET, NODEXL 등)에 관한 데이터포맷을 제공하고 매트릭스, 자카드, 유클리드언, 코사인 등의 다양한 형태값을 생성한다. 본 연구에서는 Textom의 검색어를 'OO 여행'으로 하였고 Ngram 데이터를 통해 항목간의 연결성에 대한 빈도수를 바탕으로 분석하였으며 누락된 키워드는 데이터 정제작업을 통해 보완하였다.

소셜네트워크분석(SNA)을 위한 분석도구는 'Gephi'으로 앞 단계에서 검색된 콘텐츠 간의 연결 관계를 시각적으로 표현하고 수리적인 중요성을 제시한다. 본 연구에서는 검색된 키워드 중에 주요한 단어(관광 콘텐츠, 평가 등) 간에 연결매트릭스를 작성하여 그 관계를 분석하였다.

#### 2. 조사대상

본 연구는 전라북도 지역을 중심으로 하며 구체적인 공간적 범위는 전라북도 14개 지역인 전주시(Jeonju), 군산시(Gunsan), 익산시(Iksan), 정읍시(Jeongeup), 남원시(Namwon), 김제시(Gimje), 완주군(Wanju), 진안군(Jinan), 무주군(Muju), 장수군(Jangsus), 임실군(Imsil), 순창군(Sunchang), 고창군(Gochang), 부안군(Buan) 등을 대상으로 연구를 진행하였다. 이 지역은 대부분 농·산촌지역으로 구성되어 있다.

시간적 범위는 실증 조사 데이터의 수집기간인 2015년 1월 1일부터 2015년 12월 31일까지이다. 다만, 검색 기간이 한정적인 'Social Metrics'의 경우 2016

년 4월 2일부터 2016년 9월 15일까지 약 5개월간의 SNS 및 웹사이트 데이터를 수집하였다.

### 3. 분석방법

본 연구의 방법은 국가가 제공하는 공공데이터와 웹사이트와 소셜미디어에 나타난 특성 및 여론을 분석하여 전라북도 14개 지역의 고유한 특성과 콘텐츠를 조사하고 그 결과를 바탕으로 SNA를 통해 각 지역의 콘텐츠의 분포 및 연결성을 분석한다. 소셜네트워크분석(SNA)이란 명목자료의 관계와 상호작용을 그래프로 표현하고 이를 통해 사회현상 및 구조를 파악하기 위한 방법이다(Sohn et al. 2012). 분석도구인 ‘Gephi’는 가공되지 않은 그래프데이터를 네트워크 형태로 만들어 시각화해주는 애플리케이션으로 소셜네트워크 정보를 표현하는 데 적합하다. 본 연구에서는 ‘Gephi’를 통한 분석 중 특히 Eigenvector Centrality에 따른 결과를 바탕으로 연구를 진행하였다. 네트워크분석에서 각 관광콘텐츠는 노드(node)로 표시되고 각 노드의 중요성을 표시하기 위해 여러가지 Centrality 지수(index)를 사용하는데 그 중에서 Eigenvector Centrality 값은 특정 노드의 중심성에 따른 주변 노드의 중심성을 기반으로 계산되어진다. 인접행렬(adjacency matrix)로

표현된 네트워크의 최대 고유값(eigenvalue)을 구한 후, 해당하는 고유벡터(eigenvector)를 찾아 해당 노드에 대응시켜서 구한다(Han & Lee 2015). 중심성이 큰 노드일수록 Eigenvector Centrality가 높고, 이는 네트워크상에서 가장 중심에 있고 영향력이 크다는 것을 의미한다(Han & Lee 2015). 이를 통해 전라북도의 관광지 및 관광자원과 지역별 관광지의 관광자원 간의 관계 및 특성을 파악해보았다. 본 논문에서는 14개 지역 중 풍부한 관광자원을 갖고 관광객수의 가파른 증가를 보이는 전주를 예시하고, 농촌지역으로 관광자원이 풍부한 무주, 자연자원과의 연결성이 두드러지게 나타난 순창, 관광자원 간의 연결성이 낮고 관광자원의 수가 적은 장수 등 4개 지역을 예시로 표시하였다.

## IV. 결과 및 고찰

### 1. 전라북도 관광객 현황 분석

#### 1) 전라북도 관광지 내·외국인 방문객 현황 분석

2015년을 기준으로 1월부터 12월까지 전라북도의 14개 지역을 방문한 방문객의 수요 데이터를 관광지식정보시스템(www.tour.go.kr)을 통하여 조사한 결과는 Fig. 1과 같다. 전체 29,657,053명 중 내·외국인 방문객 총계를 살펴보면 남원시가 4,686,520명으로 가장

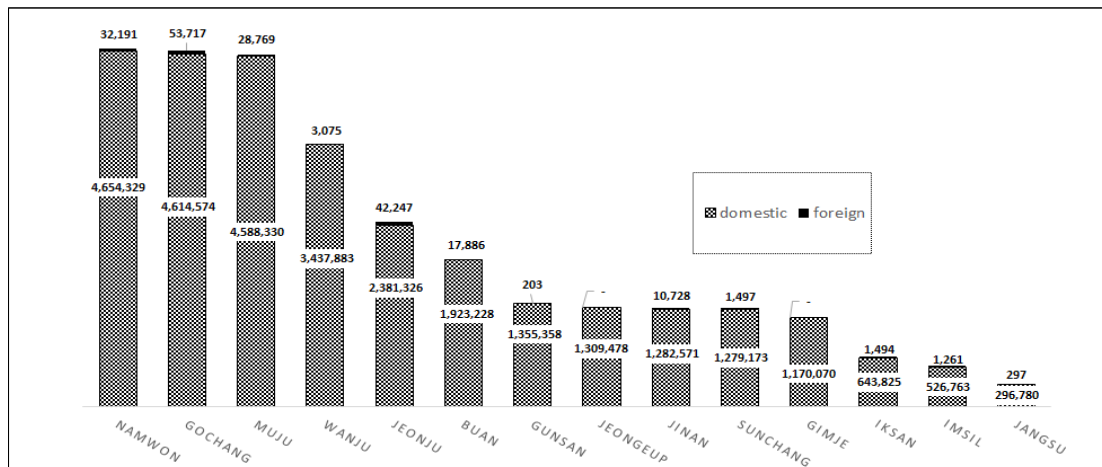


Fig. 1. Number of tourists (foreign and domestic), Jeollabuk-do, 2015  
-The Korea tourism knowledge & information system

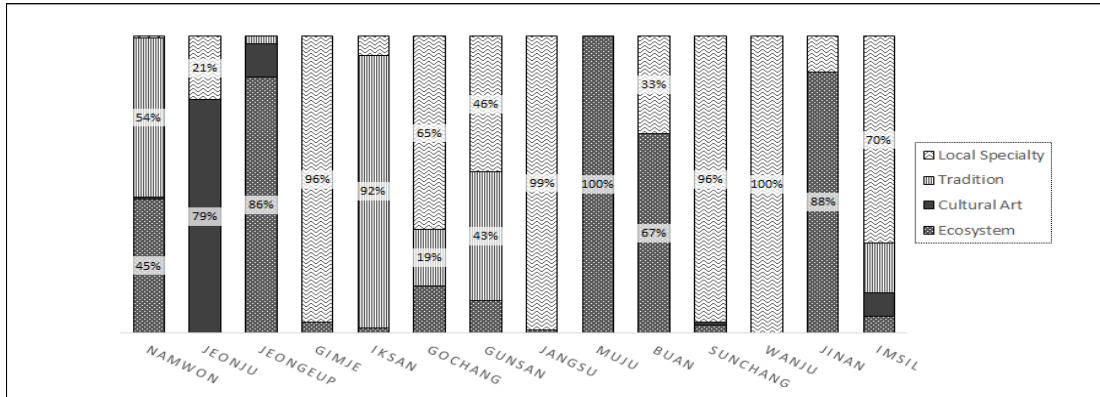


Fig. 2. Proportion of visitors in the Jeollabuk-do festival by contents, 2015

많았고 고창군은 4,668,291명, 무주군은 4,617,099명, 완주군은 3,440,958명, 전주시는 2,423,573명, 부안군은 1,941,114명, 군산시는 1,355,561명, 정읍시는 1,309,478명, 진안군은 1,293,299명, 순창군은 1,280,670명, 김제시는 1,170,070명, 익산시는 645,319명, 임실군은 528,024명, 장수군은 297,077명 순으로 나타났다. 전주보다 고창, 무주, 완주 등의 농촌지역이 관광객수가 많다는 것이 주목할 만한 사실이고, 관광객수가 적은 장수, 임실 또한 농촌지역이라는 것도 사실이다.

2) 전라북도 축제 내·외국인 방문객 현황 분석

전라북도에서는 농촌관광의 활성화를 위하여 각 지역의 특산물과 환경자원을 이용한 다양한 축제를 개최하고 있다. 그런데 이러한 축제가 관광객수의 증가에 영향을 주고 있으며 지역별 특성을 잘 반영하고 있는가를 판단하기 위해서는 축제방문객에 대한 분석이 필요하다.

2015년을 기준으로 2015년 1월부터 12월까지 전라북도 축제를 목적으로 방문한 방문객수를 14개 지역별로 전라북도 공공데이터 포털을 통해 조사한 결과는 다음과 같다. 전체 축제 내·외국인 방문객수는 6,014,966명인데, 남원시가 1,060,724명으로 가장 많았고 전주시는 706,852명, 정읍시는 704,484명, 김제시는 697,031명, 익산시는 671,440명, 고창군은 664,591명, 군산시는

350,072명, 장수군은 255,119명, 무주군은 216,365명, 부안군은 182,795명, 순창군은 181,948명, 완주군은 133,196명, 진안군은 100,852명, 임실군은 89,497명 순으로 나타났다.

이들 축제방문객을 축제의 성격에 따라 구성비율로 표시하면 Fig 2와 같다. 지역의 특산물(local specialties) 축제에 의존한 지역은 김제를 제외하고 모두 낮은 방문객수를 기록하였다. 축제관광객수가 많았던 남원, 전주, 정읍 등에는 문화예술(cultural art), 전통민속(tradition) 축제 등의 다양한 성격의 축제가 분포되었지만 한 두 종류의 축제에 치중한 지역에서는 관광객의 수가 많지 않았다는 것을 알 수 있다. 현재 각 지방자치단체가 관광객 유치를 위하여 지역특산물이나 천연자원(Ecosystem)을 이용한 축제를 개최하고 있지만, 자연자원을 활용한 축제를 여는 무주와 지역특산물 축제에 집중하는 순창, 완주, 임실 등의 축제방문객수가 적다는 사실은 한두 가지로 특성화된 축제만으로 농촌관광을 활성화하는데 한계가 있음을 보여주고 있다.

2. 전라북도 관광지 특성 네트워크 분석

빅데이터 분석 도구인 'Social Metrics'로부터 획득한 전라북도 관광지의 오피니언데이터를 분석하여 시각화한 결과는 다음 Fig 3과 같다. 각 지역별로 관광객들의 관광의견을 네트워크분석하여 각 도시들이 어떤 특성과 연관성을 갖는지를 살펴보았다.



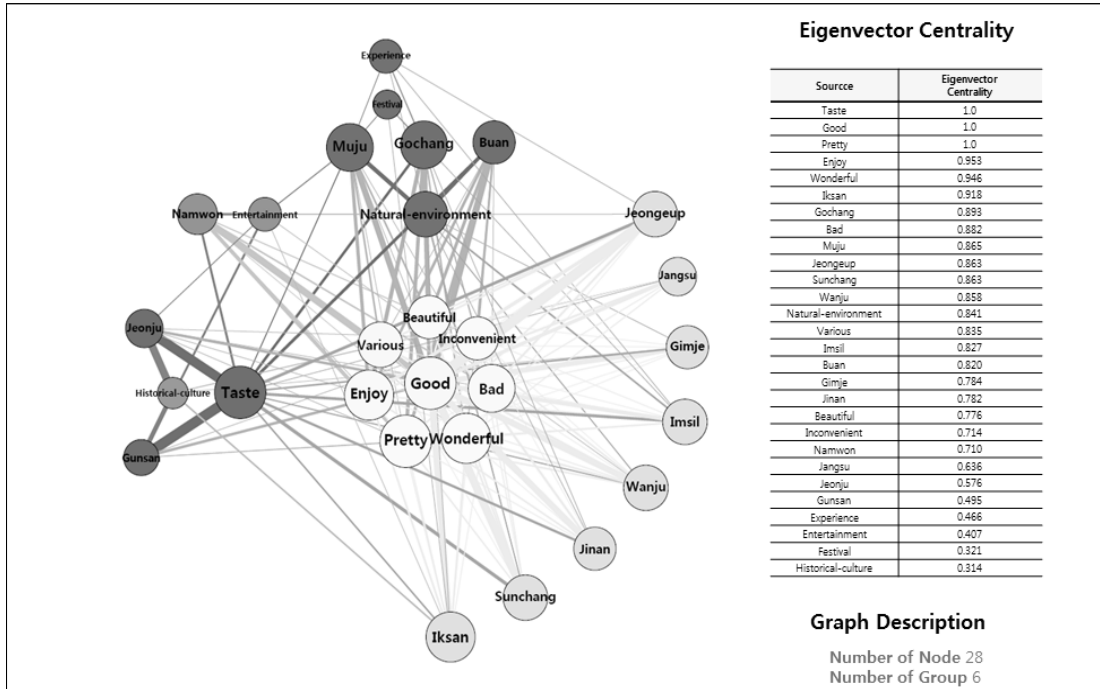


Fig. 3. Tourism characteristics in Jeollabuk-do applied social network analysis

현재 전라북도에서 관광자원으로 가장 큰 역할을 하는 콘텐츠가 맛(taste)인데, 전주와 군산은 맛, 역사 유물이라는 키워드에 의해 강하게 연결되어있다. 그리고 무주, 고창, 부안은 자연환경, 축제, 체험관광 등에 의해 강하게 연결되어 있다. 그 중간에 위치한 남원은 맛, 역사유물과 자연환경 등이 동시에 강하게 잘 연결 있으면서 좋다(good)라는 의견에도 강하게 연결되어 있어 남원이 왜 전라북도에서 가장 관광객수가 많고 관광산업이 활성화 되어있는지를 짐작하게 한다. 또한 정읍, 장수, 김제, 임실, 완주, 진안, 순창, 익산 지역에서는 각 지역을 대표하는 뚜렷한 콘텐츠가 부족하다는 것을 알 수 있다. 이러한 결과를 통해 현재 전라북도의 관광자들 중 활성화된 지역은 뚜렷한 특성과 관광 콘텐츠 간의 연결성을 갖고 있으며 그 콘텐츠는 맛, 자연환경, 문화유산, 축제, 체험, 오락 등으로 정리할 수 있다.

### 3. 지역별 관광지 특성 네트워크 분석

앞 절에서 찾아낸 특별한 관광콘텐츠의 정립과 다양성이 농촌관광을 활성화할 수 있다는 결과를 바탕으로 전라북도지역의 농촌관광 활성화 요인을 찾아보고자 각 지역별 관광 콘텐츠 간의 연결성을 네트워크분석 하였다. 이는 각 지역에서 보유하고 있는 특화된 관광자원 간에 어떤 관계를 갖는 것이 바람직한지 판단하게 해줄 것이다. 지역별관광지특성 네트워크분석은 빅데이터 분석 도구인 'Textom'으로 획득한 관광지 텍스트데이터를 활용하였다. 여기서는 전라북도 14개 지역 중 전주, 무주, 순창 장수 등의 4개 지역을 표시하였다. 농촌이 아닌데도 전주시역을 포함한 이유는 최근 들어 관광객수가 증가하고 있는 전주의 사례를 참고하고 농촌지역에서 현상과 도시지역에서의 현상에 일관성이 있음을 보여 우리의 발견한 사실이 보편성을 가짐을 입증하기 위함이다.

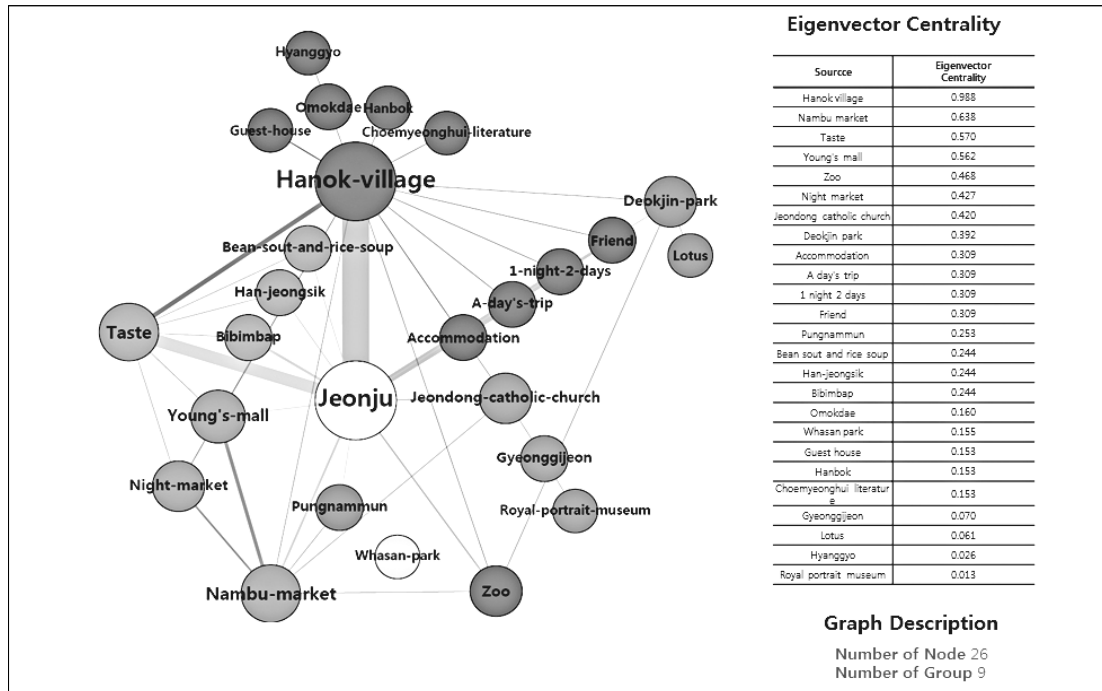


Fig. 4. Tourism characteristics in Jeonju applied social network analysis

1) 전주 관광지 특성 네트워크 분석

전주의 관광지데이터를 분석하여 시각화한 결과 Eigenvector Centrality가 높은 순으로 살펴보면 한옥 마을, 남부시장, 맛, 청년몰, 동물원, 야시장, 전통성당, 덕진공원 등의 관광요인들이 있다. Centrality가 가장 높은 노드인 한옥마을이 가장 큰 중심성을 가지고 있으며 기타 관광자원들과 밀접한 연결 관계를 가지고 있는 것을 확인할 수 있다. 한옥마을 주변의 문화역사 자원인 전통성당, 경기전과 연결되어 있고 이를 중심으로 맛 자원과도 높은 연결성을 보이고 있다. 또한 엔터테인먼트자원인 덕진공원, 동물원, 남부시장 등이 야시장과 청년몰을 포함하여 맛 자원과 조화롭게 연결된 것을 볼 수 있다. 또한 전체 연결망은 5~6개의 클러스터로 구분되는데 높은 Centrality를 갖는 콘텐츠인 한옥마을, 남부시장, 맛, 동물원, 덕진공원 등이 각 클러스터에 고루 분포되어 있는 것을 확인 할 수 있다.

2) 무주 관광지 특성 네트워크 분석

무주의 관광지 데이터를 분석하여 시각화한 결과 Eigenvector Centrality가 높은 순으로 살펴보면 덕유산, 적상산, 반디랜드, 구천동, 머루와인동굴, 무주리조트 등의 관광 요인들이 있다. Centrality가 가장 높은 노드인 덕유산이 가장 큰 중심성을 가지고 있으며 다른 관광자원들과 밀접한 관계를 가지고 있다. 자연자원인 덕유산, 적상산, 구천동이 연결되어 있고 엔터테인먼트 자원인 무주리조트를 중심으로 스노우보드, 스키장, 곤도라 등의 콘텐츠와 높은 연결성을 보이며 반딧불을 소재로 한 지역축제를 중심으로 자연자원과 조화로운 연결성을 보인다. 문화역사자원인 곤충박물관도 연결성은 낮지만 콘텐츠분포가 고르게 나타난 것을 볼 수 있다.

2015년을 기준으로 하여 전라북도 무주의 방문객수를 살펴보면 4,617,099명으로 전라북도에서 3번째로 많은 방문객을 유치하였다. 시각화한 결과 6개로 분리된 각 클러스터에 높은 Centrality를 갖는 콘텐츠가 고루

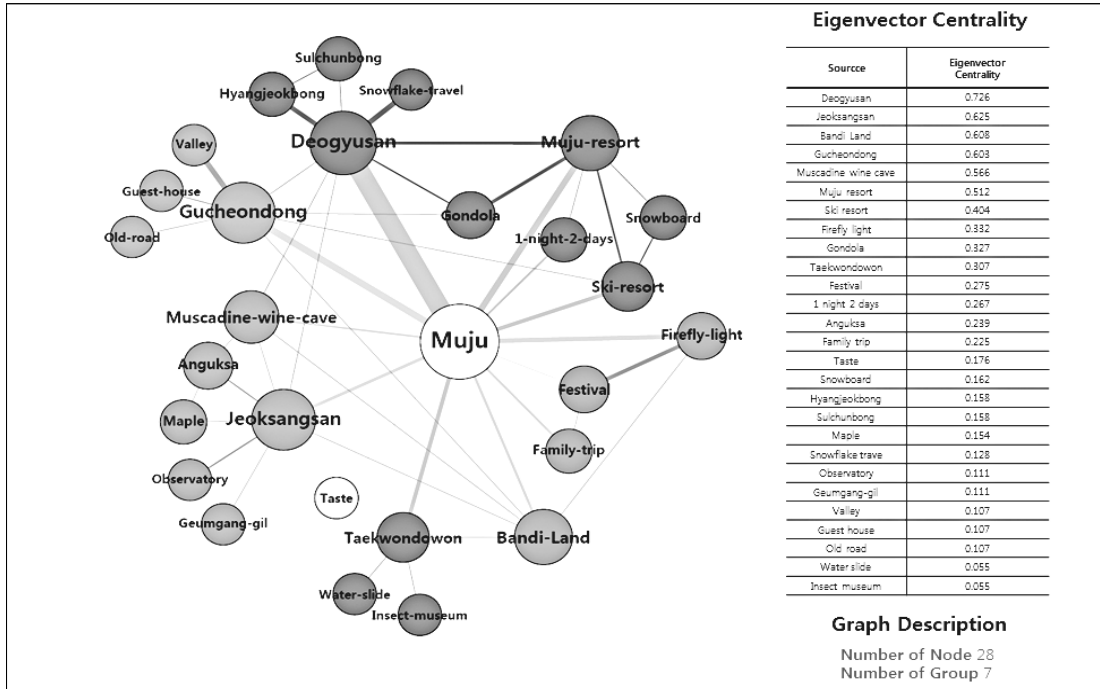


Fig. 5. Tourism characteristics in Muju applied social network analysis

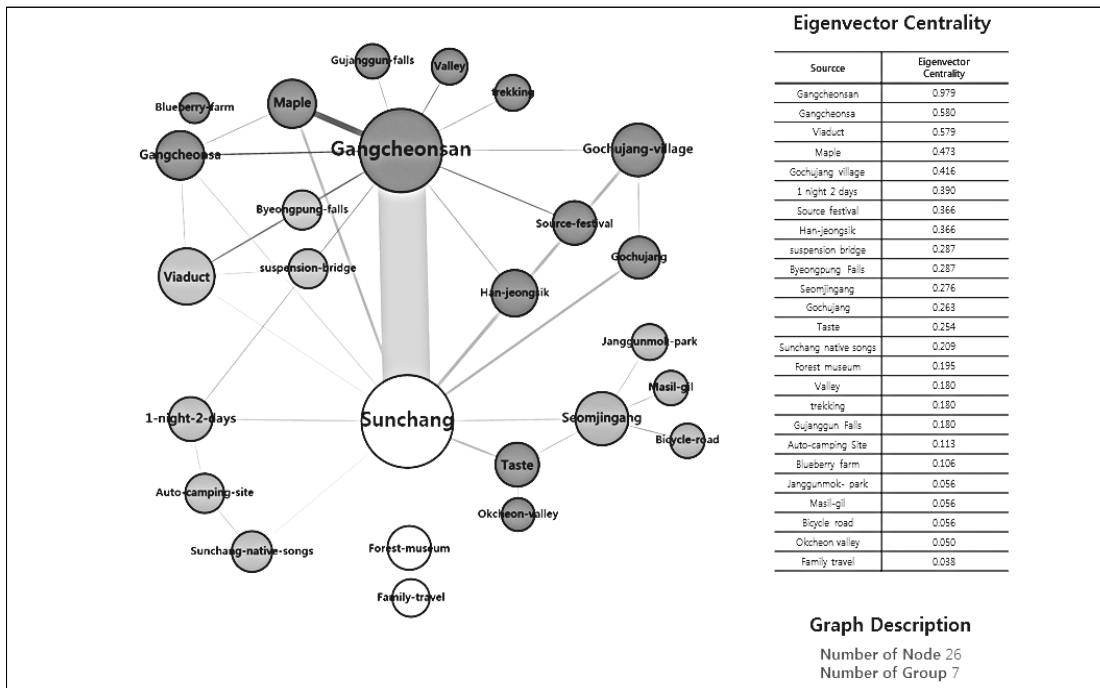


Fig. 6. Tourism characteristics in Sunchang applied social network analysis

포함되어 있고 관광자원들 간에 강한 연결성을 발견할 수 있었다. 관광자원이 풍부하고 자연환경자원, 문화역사자원, 엔터테인먼트 간의 관계가 밀접하게 연결되어 있을수록 방문객수가 많은 것을 발견할 수 있다. 전주의 네트워크 시각화자료와 비교해보았을 때, 콘텐츠 수는 무주에 비해 전주가 높은 편이나 방문객수는 무주가 현저히 높은 것으로 나타나고 있다. 이는 콘텐츠의 수보다 조화로운 연결이 더 중요한 방문객유치 요인임을 의미하고 있는데 이러한 연구를 바탕으로 콘텐츠 간의 조화로운 연결성에 대한 개발이 요구되는 시점이다.

### 3) 순창 관광지 특성 네트워크 분석

순창의 관광지 데이터를 분석하여 시각화한 결과 Eigenvector Centrality가 높은 순으로 살펴보면 강천산, 강천사, 구름다리, 단풍, 고추장마을 등의 관광 요인들이 있다. 강천산이라는 콘텐츠가 가장 높은 중심성을 가지고 있지만 다른 관광자원들과 밀접한 관계를 가지기보다는 독자적으로 높은 빈도수를 가지고 있음

을 알 수 있다. 순창은 자연자원에서 두드러진 특징을 보이는데 강천산이 병풍폭포, 구름다리 등의 자연자원과 연결되었고 순창의 특산물인 고추장과 관련한 장류축제의 개최시기인 가을과 연결시키려는 시도를 보이는 것을 알 수 있다. 그러나 그 외의 콘텐츠 간의 연결성이 낮고 콘텐츠의 수도 부족한 편이다.

2015년을 기준으로 하여 전라북도 순창의 방문객수를 살펴보면 1,280,670명으로 비교적 적은 방문객을 유치하였다. 시각화한 결과 관광자원의 수는 많은 편이었으나 관광자원들 간의 연결성이 무주에 비해 비교적 낮은 것을 알 수 있었고 빈도수가 높은 강천산을 중심으로 하여 지역의 특산물인 고추장의 연결성을 높일 필요가 있음을 알 수 있다.

### 4) 장수 관광지 특성 네트워크 분석

장수의 관광지 데이터를 분석하여 시각화한 결과 Eigenvector Centrality가 높은 순으로 살펴보면 장안산, 논개사당, 무룡고개, 의암사, 논개생가, 봉화산 등

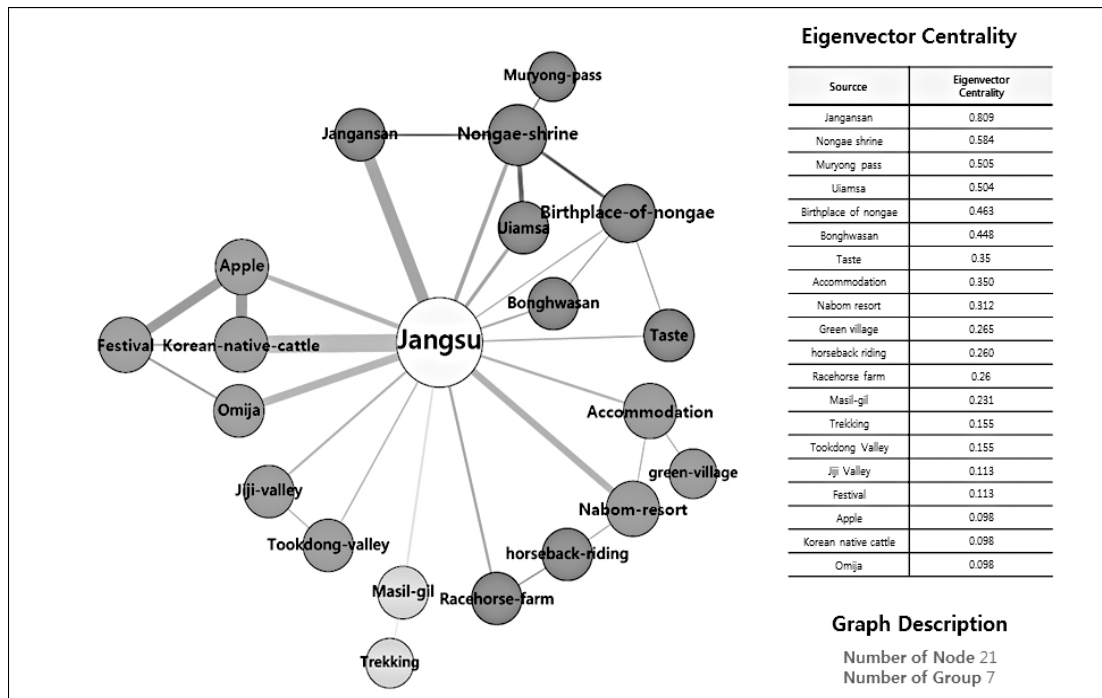


Fig. 7. Tourism characteristics in Jangsu applied social network analysis

의 관광 요인이 있다. Centrality가 가장 높은 노드인 장안산이 가장 큰 중심성을 가지고 있지만 다른 콘텐츠들과 연결성이 매우 낮은 것을 알 수 있다. 자연 자원인 장안산과 문화역사 자원인 논개사당이 연결되어 있고 미비하지만 맛집과 연결된 것도 알 수 있다. 그러나 장수지역의 각 콘텐츠들은 독립적으로 존재하고 있음을 Fig. 7에서 알 수 있다. 엔터테인먼트 자원인 경주마복장, 나뭇리조트 등도 개별적으로 존재하며 지지계곡, 토옥동 계곡인 자연환경도 다른 콘텐츠들과 연결성이 없다. 또한 순창과는 달리 지역의 특산물을 대표로 한 축제들이 다른 콘텐츠들과 연결되지 않은 것을 볼 수 있다.

2015년을 기준으로 하여 전라북도 장수의 방문객수를 살펴보면 297,077명으로 전라북도 지역 중 가장 낮은 방문객수를 유지하였다. 관광자원들의 연관성을 시각화한 결과 방문객을 유치하는 가장 큰 요인으로는 판단되는 관광자원들 간의 연결성이 매우 낮았으며 콘텐츠수도 비교적 적은 편이라는 것을 알 수 있었다. 이 사례를 통해와 연결성이 낮고 콘텐츠들이 독립적으로 존재할 경우 관광매력도가 떨어져 방문객수가 현저히 낮아진다는 사실을 알 수 있었다.

## V. 요약 및 결론

본 연구는 빅데이터 분석방법을 통하여 관광지 특성 및 관광자원간의 연결성에 따른 관광객 수요와의 연관성을 파악하고자 하였다. 이러한 연구를 달성하기 위해 관광지식정보시스템에서 제공하는 2015년 1월~12월까지 주요 관광지 및 콘텐츠에 따른 축제 관광객 데이터와 SNS를 통한 전라북도 14개 지역의 관광지 데이터를 수집하였다. 수집된 데이터를 SNA를 통하여 관광지 특성 및 관광자원 간의 연결성을 파악하였고 그에 따른 관광지 활성화 방안을 모색하였다.

본 연구의 분석 결과, 관광콘텐츠의 수와 관광객 수요는 비교적 높은 상관관계를 보였지만 관광지선택에 있어서 가장 중요한 영향을 미치는 것은 관광콘텐츠 간의 연결성인 것으로 분석되었다. 특히 무주의 경우

풍부한 관광자원과 이들의 조화로운 분포로 인해 활발한 관광활동이 이루어지고 있는 것으로 분석되었다. 순창의 경우 지역의 특성화된 콘텐츠에 비해 관광객의 수가 비교적 평이한 편이었는데 이는 문화역사, 맛 등의 자원이 부족하기 때문인 것으로 나타났다. 이러한 연구결과를 바탕으로 문제점을 보완하기 위한 콘텐츠 개발이 요구되고 있는 시점이다. 또한, 관광 콘텐츠 간의 중복성 문제도 나타났다. 유사한 특성을 가진 관광 콘텐츠의 분포는 관광객 유치를 저하시키는 것으로 분석되었다. 순창과 장수의 경우 자연자원들의 분포가 높은 편이고 이를 활용한 체험활동 등이 주로 이루어지고 있었는데 이는 다양한 콘텐츠의 부족으로 낮은 관광객 수요를 보였다고 판단되어진다. 단, 장수에 비하여 순창의 관광객의 수요는 높은 편에 속하였는데 이는 콘텐츠의 연결성이 비교적 높기 때문인 것으로 나타났다. 특히 축제와의 연결을 통한 관광 콘텐츠를 연결 짓는 시도를 하고 있는 것으로 보였다.

이러한 분석 결과를 바탕으로 농촌관광은 본래부터 가지고 있는 천혜의 자연과 문화역사, 맛, 엔터테인먼트, 축제, 체험 등으로 콘텐츠가 이루어져있고 이들 간의 연결성이 높을수록 방문객수가 높다는 것을 알 수 있었다. 그러므로 농촌관광 활성화를 위해서는 관광콘텐츠의 적절한 개발과 관광콘텐츠 간의 조화로운 연결성이 필요하다고 할 수 있다.

관광콘텐츠 개발을 위해 다음과 같은 점이 고려되어야 한다. 첫째, 대부분의 지역에서 큰 비중을 차지하고 있는 자연자원을 중심으로 하여 각 콘텐츠들이 골고루 분포할 수 있는 방안을 생각해야 한다. 둘째, 콘텐츠 간의 환경, 콘텐츠 간의 연결 시기, 콘텐츠의 집중성 제거 등의 요인을 파악하여 콘텐츠 연결성에 기반을 둔 발전전략을 수립해야 한다. 또한 각 지역별 방문객의 기대수준을 파악하고 그 결과를 토대로 관광 콘텐츠를 개발해야 하는데, 즉 각 지역의 콘텐츠뿐만 아니라 방문객에 대한 연구를 병행함으로써 방문객에 대한 깊은 이해를 바탕으로 상품개발이 이루어져야 할 것이다.

전라북도 지역별 관광자원 연결성 증대방안에 대해

살펴보면 첫째, 콘텐츠를 조화롭게 연결시키기 위해서는 주변환경을 고려해야한다. 콘텐츠 간의 거리 및 교통 등으로 접근성이 용이해야 관광객의 접근도 원활하게 이루어질 것이고 이는 관광객유치에 긍정적인 영향을 끼칠 것이다. 두번째로 콘텐츠 간의 연결시기를 고려한 콘텐츠개발이 이루어져야 한다. 콘텐츠가 가진 속성을 잘 파악하고 자연자원의 계절성, 축제 개최시기 등을 고려한 연결과 그에 따른 콘텐츠의 개발은 관광지의 시기별 조화를 통해 관광객의 관광활동을 촉진시킬 것이다. 세번째로는 콘텐츠 간의 중복성 제거를 통해 콘텐츠의 다양성을 높여야 한다. 농촌관광의 콘텐츠의 품질 저하 및 중복으로 인해 흥미유발을 하지 못함으로써 농촌관광의 유인 요인이 취약한 실정이다. 이러한 문제점을 극복하기 위해서는 분석결과를 바탕으로 중복된 콘텐츠들을 보완하여 콘텐츠의 특성을 다양화하여 관광객을 유치할 수 있어야 한다. 주어진 자연자원에 집중하기보다 연결이 용이한 문화역사, 먹거리, 엔터테인먼트 등의 콘텐츠를 개발하여야 한다.

본 연구의 한계점으로는 관광과 관련된 빅데이터의 수집기간이 짧았다는 점이다. 향후 연차적으로 계절별로 자료를 수집하여 좀 더 심층적인 분석이 필요하다고 할 수 있다. 따라서 관광 빅데이터를 꾸준히 수집하고 축적된 자료를 제공할 수 있는 빅데이터 전문 수집기관이 필요할 것으로 보인다.

## References

- Agricultural Research & Extension Services(2010) Development of manual for area network on Jeonbuk rural tourism experience
- Cho WS, Cho A, Kwon KE(2015) Implementation of smart chungbuk tourism based on SNS data analysis. *J Korean Data Info Sci Soc* 26(2), 409-418
- Duan W, Cao Q, Yu Y, Levy S(2013) Mining online user-generated content: using sentiment analysis technique to study hotel service quality. 47th Hawaii International Conference on System Sciences, 3119-3128
- Han CG, Lee SH(2015) Analysis of effect of an additional edge on eigenvector centrality of graph. *J Internet Comput Serv* 16(5), 39-47
- Jeon MJ(2015) A study on the pattern of using services and the characteristics by making use of big data: a case applied to the korail service package 'Railro'. MS thesis, Graduate School Myongji University
- Jeong M, Jeon MM(2008) Customer reviews of hotel experiences through Consumer Generated Media(CGM). *J Hosp Lei Mark* 17(1-2), 121-138
- Jung KS(2015) Travel SNS recommend using the LDA-based data modeling techniques. MS thesis, Graduate School Chonnam University
- Kim HC, Kim MJ, Shin HJ(2014) Method for analysis of social data service design and implementation of Jeju tourism trends. *e-Biz S* 15(3), 173-193
- Kim NW(2015) The study on the relationship between tourism infrastructure and regional tourist demand by applying big data analysis. MS thesis, Graduate School Hanyang University
- Kim YB(2016) An importance-performance analysis of hotel selection attributes by using big data: comparison of deluxe and mid-price hotel. MS thesis, Graduate School Catholic University
- Kong HS, Song EJ(2013) A study on hotel CRM(Customer Relationship Management) using big data and security. *Conv Secu J* 13(4), 69-75
- Lau KN, Lee KH, Ho Y(2005) Text mining for the hotel industry. *J Cornell University* 34(3), 344-362
- Lee HA, Law R, Murphy J(2011) Helpful reviewers in tripadvisor, and online travel community. *J Travel Tour Mark* 28(7), 675-688
- Lee YJ, Jeong HK, Boo SH(2010) Evaluation of the publicity on etourism information: focused on the portal site 'N'. *J Tour Leis Res* 22(3), 389-408
- Lee YJ, Yoon JH(2014) A study on utilizing SNS big data in the tourism studies: based on an analysis of key words for tourism information search. *Int J Tour Hosp Res* 28(3), 5-14
- Li G, Law R, Vu HQ, Rong J, Zhao XR(2015) Identifying emerging hotel preferences using emerging pattern mining technique. *Tour Manag* 46(1), 311-321
- O'Connor P(2010) Managing a Hotel's image on tripadvisor. *J Hosp Mark Manag* 19(7), 754-722
- Oh IK, Lee TS, Chon CN(2015) A study on awareness of Korea tourism through big data analysis. *J Tour Sci* 10(136), 107-126
- Park JP(2015) A study on the activation methods through the investigation of the rural experience tourism

- village—a case of integrated Cheongju City—. *J Korean inst Rural Arch* 17(2) 17–26
- Park SH(2013) Development of the Rural tourism for rural 6th industrialization. *Agricultural Policy Focus*(Korea Rural Economic Institute) 66, 1–21
- Schuckert M, Liu X, Law R(2015) Insight into suspicious online rating: direct evidence from tripadvisor. *Asia Pacific J Tour Res* 21, 259–272
- Sim YS(2016) A study on the determinants of destination image using text mining. MS thesis, Graduate School Sejong University
- Sohn JS, Cho SW, Kwon KL(2012) Improved social network analysis method in SNS. *J Intell Inform Sys* 18(4), 117–127
- Sung KM(2003) An activation plan of jeollabuk-do regional tourism. *Proceedings 2003 Summer Symposium of the Korean Association for Public Administration*, 121–137
- Xiang Z, Schwartz Z, Gerdes JH, Uysala M(2015) What can big data and text analytics tell us about hotel guest experience and satisfaction? *Inter J Hosp Manag* 44, 120–130
- Ye Q, Law R, Gu B(2009) The impact of online user reviews on hotel room sales. *Inter J Hosp Manag* 28, 180–182
- Yoo JD(2014) A platform design for a tourism SNS issue analysis in real time using big data. MS thesis, Graduate School Kwandong University
- Yoon YS, Park DB, Lee MS(2009) Segmentation and characteristic analysis of rural tourists behavior applied by two stage cluster and multi nominal logit model. *Korean J Tour Res* 23(4), 337–355