

SNS 데이터 분석을 통한 스마트 충북관광 구축[†]

조완섭¹ · 조아² · 권가은³ · 류관희⁴

¹충북대학교 경영정보학과 · ^{2,3}충북대학교 비즈니스데이터융합학과 · ⁴충북대학교 소프트웨어학과

접수 2015년 2월 23일, 수정 2015년 3월 19일, 게재확정 2015년 3월 24일

요약

SNS 등 온라인에 쏟아지는 고객의 감성분석을 활용하여 마케팅 등 다양한 분야에서 성과를 내고 있다. 고객의 제품사용 후기 데이터를 수집하여 신제품의 마케팅 전략과 제품개선 방안을 도출하고 있으며, 선거철에는 유권자들의 마음을 실시간으로 읽어서 선거 전략을 수립하고 있다. 본 연구에서는 온라인 상에서 충북관광에 관한 글들을 수집하여 텍스트 마이닝 기법을 활용하여 분석함으로써 충북관광을 위한 홍보전략과 관광진흥 정책을 수립하는 방안을 제시한다. 사용된 데이터는 주요 블로그, 뉴스, 관광관련 공공 데이터 등이며, 충북전역의 관광지를 대상으로 한다. 분석결과 충청북도 북부지역의 관광지들이 주요 이슈 관광지로 나타났으며, 관광지별로 연관 관광지와 연관 키워드들이 분석되었다. 또한, 주요 관광지별 감성분석을 통하여 긍/부정 현황과 그 원인을 파악할 수 있도록 하였다. 특히 이들 분석에 다차원 분석 기법을 연계함으로써 관광관련 텍스트 정보를 다각도로 분석하여 전략을 수립하는데 활용할 수 있도록 하였다.

주요용어: 감성분석, 다차원 분석, 빅데이터 분석, 스마트 관광, 연관분석, 이슈분석, 충북관광.

1. 머리말

최근 정보통신기술의 발달에 따라 빅데이터 기술이 공공기관에서도 다양하게 활용되고 있다. 행정서비스 답답 또는 관행에 따른 정책수립에서 탈피하여 민간 및 공공기관이 보유하고 있는 다양한 정형·비정형의 빅데이터를 분석하여 과학적인 정책 수요조사 및 관광정책을 발굴·수립할 수 있으며, 민간·공공 데이터의 융·복합 분석을 통한 효율성 및 가능성을 도출할 수 있기 때문에 이러한 추세는 앞으로 더욱 강조될 것으로 보인다. 국민권익위원회, 한국관광공사 등에서 공공 데이터를 통합 분석하여 통찰력을 발견하는 것이 그 예이다.

민간에서도 SNS 등 온라인에 쏟아지는 고객의 감성분석을 활용하여 마케팅 등 다양한 분야에서 성과를 내고 있다. 고객의 제품사용 후기 데이터를 수집하여 신제품의 마케팅 전략과 제품개선 방안을 도출하고 있으며, 선거철에는 유권자들의 마음을 실시간으로 읽어서 선거 전략을 수립하고 있다.

본 논문에서는 온라인상에서 충북관광에 관한 글들을 수집하여 텍스트 마이닝 기법을 활용하여 분석함으로써 충북관광을 위한 홍보 전략과 관광 진흥 정책을 수립하는 방안을 제시한다. 사용된 데이터는 주요 블로그, 뉴스, 관광관련 공공 데이터 등이며, 충북 전역의 관광지를 대상으로 한다. 특히 다차원 분

[†] 이 논문은 2012학년도 충북대학교 학술연구지원사업의 연구비 지원에 의하여 연구되었음.

¹ 교신저자: (362-763) 충북 청주시 서원구 충대로 1, 충북대학교 경영정보학과, 교수.

E-mail: wscho@chungbuk.ac.kr

² (362-763) 충북 청주시 서원구 충대로 1, 충북대학교 비즈니스데이터융합학과, 석사과정.

³ (362-763) 충북 청주시 서원구 충대로 1, 충북대학교 비즈니스데이터융합학과, 석사과정.

⁴ (362-763) 충북 청주시 서원구 충대로 1, 충북대학교 소프트웨어학과, 교수.

석 기법을 연계함으로써 관광관련 텍스트 정보를 다각도로 분석하여 전략을 수립하는데 활용할 수 있도록 한다.

본 논문의 공헌은 다음과 같다. 먼저, 정보화 사회의 트렌드에 맞는 과학적 행정을 추진하여 도민으로부터 도정의 신뢰성을 제고할 수 있는 기반을 마련할 수 있다. 다음으로, 정부 3.0의 중점추진과제 중에 하나인 “빅데이터를 통한 과학적 행정 구현”에 걸맞는 관광행정을 추진한다는 점이다. 또한, 충북을 방문하는 관광객이 원하는 맞춤형 행정서비스 실현할 수 있다는 점이다. 이러한 공헌은 온라인상에 흩어져 있는 충북관광에 관한 데이터를 수집하여 통합하고, 이를 다차원적으로 분석함으로써 유용한 통찰력을 발견하는 빅데이터 기술을 기반으로 하고 있다.

논문의 구조는 다음과 같다. 제 2절에서는 관련연구를 기술한다. 제 3절에서는 연구내용 및 방법을 소개한다. 제 4절에서는 관광 데이터 분석결과를 소개한다. 이슈분석, 연관분석, 감성분석 기법으로 구성된다. 마지막으로 제 5절에서 정책적 시사점과 결론을 기술한다.

2. 관련연구

본 장에서는 관광산업에 ICT (Information and Communication Technology) 혹은 빅데이터 기술을 접목한 연구들을 소개한다. 이러한 연구는 아직 연구논문으로 발표된 것은 많지 않으며, 관광공사 등 일부 기업의 연구보고서로 발표되고 있는 상황이다.

Moon (2013)은 지능화·고도화된 ICT를 활용한 관광서비스 정책과 현황을 살펴보고, 관련 국내의 사례 분석을 통해 경제성장과 국민의 삶의 질 향상을 가져올 수 있는 ICT 기반 관광서비스의 발전방안을 모색해 보았다. 특히, 관광 서비스의 고도화 및 지능화를 위해 데이터 분석기반의 맞춤형 서비스 제공이 필요함을 강조하였다. Lee와 Yoon (2014)은 관광분야 활성화를 위한 SNS 빅데이터 분석 활용 방법을 모색하였다. 또한, 관광 데이터마이닝 대회를 통하여 관광산업 발전 정책과 마케팅 방안을 도출하는 시도도 있었다.

빅데이터의 미래 추세 예측과 마케팅 활용도가 높음에도 불구하고 아직 관광학계나 관광산업 분야에 빅데이터를 실제로 적용한 사례는 극히 드물다. 빅데이터를 활용한 관광 관련 연구로는 외식이나 호텔 등 제한된 분야로 한정하여 연구되고 있으며 (Kong과 Song, 2013; Do 등, 2013; Do와 Ryu, 2013). 온라인 관광 정보의 공공성 분석을 위해 온라인 빅데이터를 활용한 Lee 등 (2010)의 연구 등 몇몇 연구에 그치고 있어 타 분야에 비하여 빅데이터의 활용도가 낮다.

최근 들어 정부 3.0을 계기로 공공 빅데이터 개방이 시작되었고, 빅데이터 기술도 점차 안정적으로 확산되고 있어 향후 빅데이터를 활용한 관광 활성화에 연구가 활발하게 진행될 것으로 판단된다. 개방되는 공공데이터의 질적인 차원의 문제는 여전히 해결되어야 할 과제로 지적되고 있으나 이제 수집 가능한 수많은 자료를 어떻게 의미 있는 정보로 전환할지에 대한 고민이 필요한 시점이다.

3. 연구내용과 방법

본 장에서는 연구내용과 방법 및 연구 범위를 설명한 다음 본 연구의 프로토타입을 구현하기 위한 주요 패키지를 소개한다.

3.1. 연구내용

본 논문의 주요 연구내용은 다음과 같다. 먼저 온라인상의 이슈 분석을 통해 충북의 각 지역별 특화된 관광자원 개발 정책 방향 제시한다. 이슈 분석을 위해 데이터 웨어하우스 구축을 통한 다차원 분석을 수행하고, 전체/지역/테마/채널별로 상위 빈도를 차지하는 관광명소와 이슈를 분석해 낸다. 충북 관광자원에 대한 도·내의 사람들의 온라인상의 관심사를 분석하여 이슈관광 명소를 계절별로, 지역별로 도출해 낸다.

다음으로 관광 키워드들 사이의 연관분석을 통하여 패키지 관광 등을 기획하는데 근거자료를 제시하도록 한다. 어떤 관광지들이 함께 언급되는지? 어떤 관광지와 어떤 키워드 (예, 식당이나 인접 관광지 등)가 연관성이 높은지를 파악하여 패키지로 묶는 관광 상품을 개발하는데 유용할 것이다.

또한, 수집된 온라인 글을 대상으로 온라인상에서 관광객의 선호 및 불만사항을 분석하여 충북 관광 자원의 서비스 개선 방안 제시한다. 충북도내에는 350여개의 관광지가 있으며, 관광지별로, 계절별로 긍/부정으로 구분하여 제시함으로써 개선을 유도한다.

마지막으로 관광 빅데이터를 저장하고 관리할 수 있는 확장성 있는 방안을 제시하고, 적절한 범위의 데이터를 수집, 저장, 분석하는 프로토타입 시스템을 구축한다. 그리고, 분석결과가 도청 및 산하기관의 관광관련 부서의 주요 업무와 연계되어 시너지 효과를 창출할 수 있도록 하기 위함이다. 특히, 가시화를 통하여 분석결과가 의미하는 바를 직관적으로 표출할 수 있도록 한다.

3.2. 연구방법

Figure 3.1은 본 연구의 전체 프로세스를 보여준다. 본 연구는 자체 웹 크롤러를 이용하여 비정형 데이터 (뉴스, 블로그)를 수집하고, 이를 공공 데이터와 통합하여 충북 관광의 이슈, 연관 키워드, 감성분석을 수행한다. 분석결과의 시각화를 통하여 직관적인 의사결정을 지원하도록 한다.

수집 대상이 되는 정형 데이터 소스는 다음과 같다.

- 최근 5년간 지자체별 관광 만족도 보고서 (충청북도 도청)
- 충북지역 관광 명소명 DB 및 관련 정보 (충북 나드리 홈페이지)
- 주요 관광지별 입장객 통계 (관광지식 정보시스템)

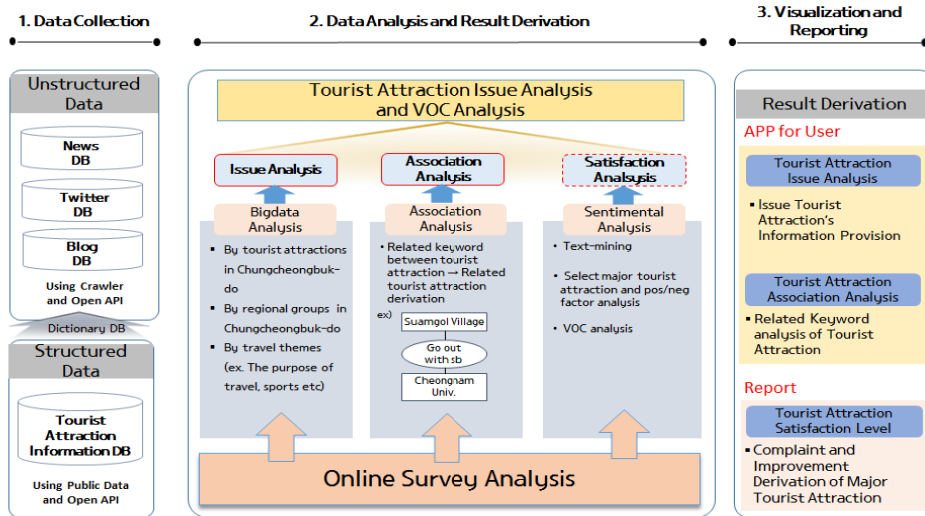


Figure 3.1 Total process of the research

다음으로 비정형 데이터는 자체 웹 크롤러를 구축하여 수집하였다.

- 데이터 수집 기준 : 충북 나드리 홈페이지에서 수집된 관광명소 키워드를 기준으로 하여 뉴스 및 블로그 게시글과 빈도수 (건수) 수집
- 데이터 수집 기간과 양 : 2013.9.01.~2014.8.31., 1,104,054건
- 데이터 출처 : 네이버 (블로그), 신문사 (뉴스), 구글 (전체)

본 연구에서 사용한 분석도구는 다음과 같다.

- 이슈분석 : MS SQL 2012 DBMS, Analysis Service
(데이터베이스 관리 시스템 및 다차원 분석 툴)
- 연관분석 : Neo 4J (그래프 데이터베이스)
- 감성분석 : 자체개발 텍스트 마이너 (마이닝 도구)

4. 분석 결과

본 장에서는 다차원 분석 기반의 이슈분석, 연관분석, 감성분석의 결과를 소개하고, 프로토타입 시스템의 구축에 관해 설명한다.

4.1. 이슈분석

블로그 및 뉴스에서 충북 관광명소 (381개)를 대상으로, 언급 빈도를 분석한 결과는 다음 Table 4.1과 같이 나타났다. 빈도 분석 결과, 뉴스 또는 블로그에서 언급이 많은 관광명소는 충주호 (Chungjuho Lake), 제천 의림지 (Jecheon Uirimji), 제천 청풍호(Jecheon Cheongpung Lake) 순이며, 10개의 상위 관광명소 중에 자연 및 휴양 관광자원이 절반 이상을 차지하는 것으로 보아, 충북관광에서 자연 관광명소에 관한 언급이 많다는 것을 알 수 있다.

Table 4.1 Top-10 Issue analysis of Chungbuk

Rank	Tourist Attraction	Frequency	Theme
1	Chungjuho Lake	15232	Nature
2	Jecheon Uirimji	7887	Nature
3	Jecheon Cheongpung Lake	5197	Nature
4	Chungju Suanbo	5145	Healing
5	Danyang Dodamsambong	4453	Nature
6	Jecheon Bakdaljae Pass	3916	Nature
7	Chungju Suamgol Village	3610	Experience
8	Chungju Zoo	3124	Healing
9	Danyang Gosu Cave	2868	Nature
10	Goesan Sanmagiyet-gil	2796	Leports

이슈 관광 명소는 지역별 (시군별), 권역별 (북부, 중부, 남부)로 구분하여 다차원 분석을 할 수 있다. Figure 4.1은 계절별로 다차원 분석한 이슈 관광명소를 보여주고 있다.

Spring		Summer		Fall		Winter	
Tourist Attraction	Frequency	Tourist Attraction	Frequency	Tourist Attraction	Frequency	Tourist Attraction	Frequency
1 Chungjuho Lake	4243	1 Chungjuho Lake	4189	1 Chungjuho Lake	3820	1 Chungjuho Lake	2980
2 Jecheon Uirimji Reservoir	2210	2 Jecheon Uirimji Reservoir	2803	2 Jecheon Cheongpung Lake	1109	2 Jecheon Uirimji Reservoir	2030
3 Jecheon Cheongpung Lake	1692	3 Jecheon Bakdaljae	2726	3 Danyang Dodamsambong Peaks	1075	3 Chungju Suanbo	1090
4 Chungju Suanbo	1154	4 Chungju Suanbo	1966	4 Chungju Suanbo	935	4 Danyang Dodamsambong Peaks	897
5 Chungju Suamgol Village	1026	5 Jecheon Cheongpung Lake	1685	5 Jecheon Uirimji Reservoir	844	5 Cheongju Zoo	764
6 Danyang Dodamsambong Peaks	990	6 Danyang Dodamsambong Peaks	1491	6 Chungju Suamgol Village	738	6 Jecheon Cheongpung Lake	711
7 Goesan Sanmagi old roads	910	7 Chungju Suamgol Village	1213	7 Cheongju Zoo	713	7 Chungju Suamgol Village	633
8 Chungju Sangdangsanseong Fortress	727	8 Danyang Gosu Cave	1158	8 Goesan Sanmagi old roads	665	8 Chungju Sangdangsanseong Fortress	521
9 Danyang Gosu Cave	646	9 Cheongju Zoo	1117	9 Danyang Gosu Cave	658	9 Yeongdong Minjujisan	456
10 Danyang Guinsa Temple	597	10 Goesan Sanmagi old roads	841	10 Danyang Guinsa Temple	483	10 Eumseong Great Stone Face Sculpture Park	436

Figure 4.1 Analysis result of issue tourist attractions based on season

4.2. 연관분석

연관분석은 관광 명소별 연관성 분석, 관광명소-키워드 연관성 분석, 다차원 분석과 결합한 연관분석 등 다양한 형태로 분석될 수 있다. Figure 4.2는 관광명소들 간의 연관성 분석 결과의 일부를 보여주고 있다. Figure 4.2의 노드는 각각 텍스트 (관광명소)를, 링크 위의 숫자는 노드 (관광명소) 간의 연결강도를 의미한다. 충주호 (Chungjuho Lake)는 수안보 (Chungju Suanbo) 및 충주호 유람선 (Chungjuho Lake cruise ship)과 연관성이 있으며, 이들을 패키지로 묶어서 관광 상품화 할 필요가 있음을 의미한다.

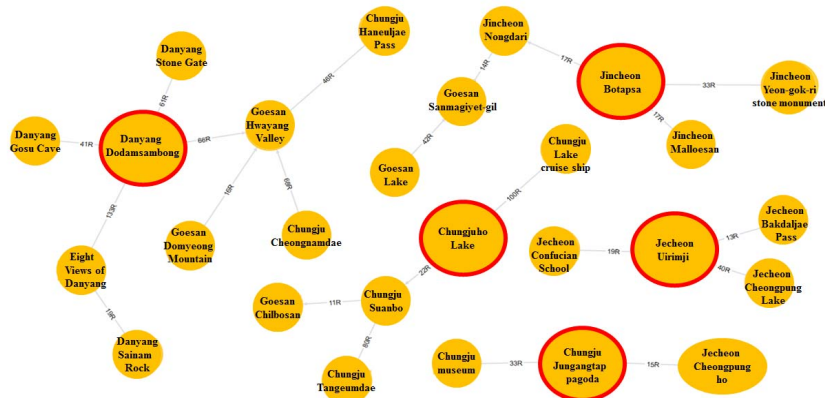


Figure 4.2 Association analysis of tourist attractions in Chungbuk

Table 4.2는 관광명소별 연관분석 결과 강도가 높은 순서대로 표시한 것이다. 이는 Figure 4.2와 같이 그래프 형태로 시각화할 수도 있다.

Table 4.2 Association analysis of attractions based on season

	Association tourist attraction	Keyword	Count
Danyang Dodamsambong	Danyang Eight Sceneries	famous, cruise ship, landscape, Oksunbong Peak, Gudambong Peak	133
	Goesan Hwayang Valley	famous	66
	Danyang Stone Gate	-	61
	Danyang Gosu Cave	famous, filming, cruise ship, cool, nature, family.kid	41
Chungjuho Lake	Chungjuho Lake cruise ship	-	100
	Chungju Suanbo	-	22
Jincheon Botapsa	Jincheon Yeon-gok-ri stone monument	-	33
	Jincheon Malloesan	-	17
	Jincheon Nongdari	heart	17
Jecheon Uirimji	Jecheon Cheongpungho	tree, wind, landscape	40
	Jecheon Confucian School	-	19
	Jecheon Bakdaljae Pass	nature, morning, landscape	13
Chungju Jungangtap pagoda	Chungju museum	culture,history	33
	Chungju Seven-storied stone pagoda in Tap-pyeong-ri	-	15

본 연구에서는 연관성 분석을 다차원 분석과 결합하여 더욱 세분화된 분석결과를 도출하도록 한다. Figure 4.3은 계절별 (봄)로 관광지별 연관성을 분석한 결과를 보여주고 있다.

이는 Figure 4.2의 전체 연관성 분석결과와는 다른 양상을 보여주며, Cho (2014)에서 보고한 바와 같이 여름, 가을, 겨울의 연관성 분석결과는 각각 그 특색을 가지고 있다. 이러한 다차원 분석은 계절별 뿐 아니라 날씨별, 년도별, 심지어 엑스포 등 이벤트별로 다양한 각도에서 분석될 수 있다.

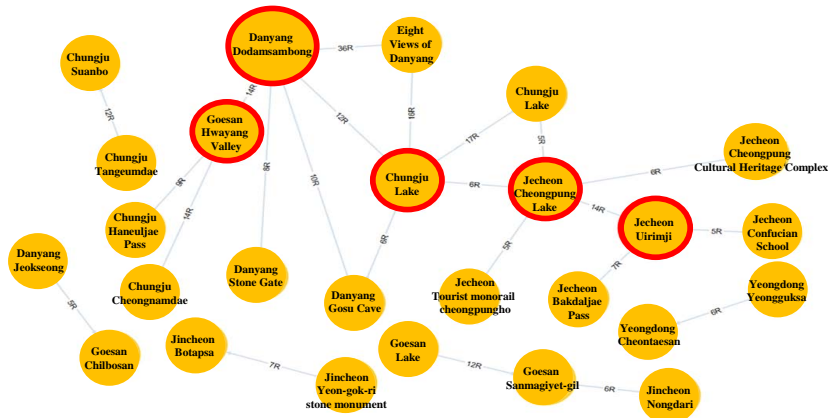


Figure 4.3 Association analysis of tourist attractions in spring

4.3. 감성분석

감성 분석은 관광지별로 고객들의 긍정과 부정적인 의견을 수집하여 분석하는 것으로 긍정 및 부정의 빈도수와 함께 그 원인을 파악할 수 있으므로 관광정책을 수립하는데 중요한 단서를 제공할 수 있다.

본 논문에서는 충청북도 주요 관광지에 대한 긍/부정 의견을 온라인에서 수집하여 분석하였다. Figure 4.4는 청주시 청남대에 대한 긍/부정 의견을 분석하여 시간의 흐름과 주요 주제별 긍정과 부정의 빈도수를 시각화한 것이다. Figure 4.4의 X축은 카테고리별로 시간의 흐름 (월 단위)을 나타내고, Y축은 긍정 (파란색)과 부정 (주황색)의 빈도수를 나타낸다. 왼쪽부분의 그래프는“관광교통 및 정보”측면에서 시간의 흐름에 따른 긍/부정 의견의 추이를 나타내고, 그 다음은 “관광시설 및 안내시설” 측면에서 긍/부정의 추이를 보여주고 있다. 전체적으로 충주호는 오른쪽 부분에서 자연경관 및 관광명소 측면에서 만족도가 매우 높은 관광지임을 나타내고 있다.

Table 4.3은 충주호 (Chungjuho Lake)에 대하여 주제별로 분기별로 긍정과 부정의 상세정보를 분석한 결과이다. 여기서는 지면관계상 충청북도의 주요 관광 교통 및 정보 주제에 한하여 분석한 결과만 수록하고, 나머지는 Cho (2014)를 참고하기 바란다.

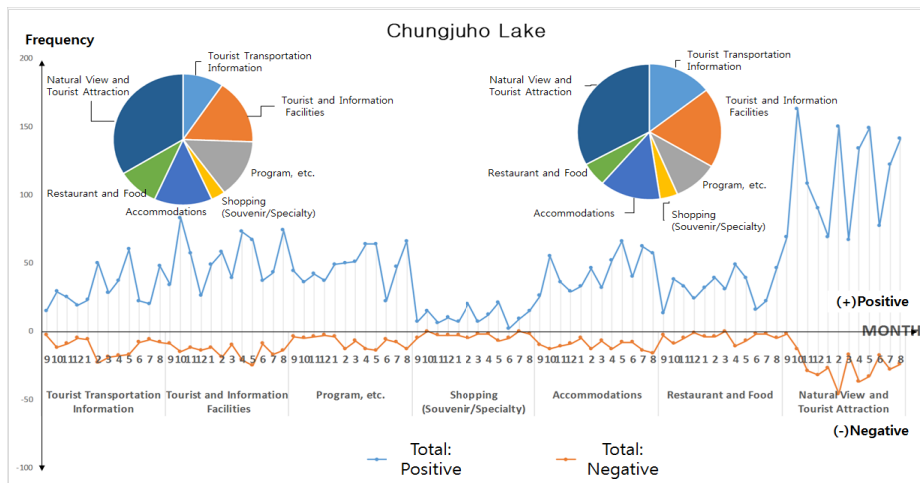


Figure 4.4 Positive/negative frequency analysis of Chungjuho lake based on categories

의 필요성이 제기되고 있다. 아울러 대청호 주변 개발 제한 구역을 청정관광구역으로 개발함으로써 수자원 보호와 관광 증진을 동시에 해결할 필요가 있음을 알 수 있었다. 북부권의 경우 충북의 대표 관광 명소인 충주호를 중심으로 인접 관광명소를 연계 육성하고, 남부권의 경우 온라인 홍보 등을 통하여 자연경관 등 대표관광명소를 홍보하는 것이 필요한 것으로 나타났다. 연관 키워드 및 연관 관광명소를 분석한 결과 제천·단양은 당일 코스로 여러 관광명소들을 방문하는 것으로 나타났으며, 충주는 1박 2일 코스로 충주호 주변의 캠핑장을 이용하며 충주호, 유람선 등 충주권역 내에서만 관광을 하는 것으로 나타났다 (Figure 4.3). 따라서 북부권의 제천, 단양, 충주를 연계하여 인접 지역 내의 관광명소들을 연계한 관광 상품 개발이 필요한 것으로 보인다.

도내 11개의 관광명소 감성분석 결과 공통적으로 나타난 불만사항은 대중교통을 이용하는 사람들이 관광명소 입구까지 가는데 시간이 오래 걸려 타 지역의 다른 관광명소를 택하는 경우가 있었으므로 관광명소의 교통 및 서비스 개선이 필요하다 (예: 자전거 등의 새로운 관광수단 도입). 향후에 교통 데이터 (버스 혹은 택시) 분석을 추가하여 주요 관광지 버스배차 간격 개선과 노선조정 등에 활용할 수 있을 것이다. 또한 충북 관광명소의 경우 각 관광명소의 특색을 나타내는 기념품이 없는 상황이므로 충북의 관광 수입을 증가시키기 위한 충북 관광명소만의 특색 있는 기념품 개발 등 대안이 제시되어야 할 것이다.

향후 과제로는 통신사 등의 이동정보 데이터를 추가 활용하여 관광객 유입경로와 충북내 이동경로 분석, 카드사용정보 등과 연계한 관광수익성 분석 등이 통합 분석되면 더욱 효과적이고 세밀한 관광정책 수립이 가능해질 것이다. 가장 중요한 것은 본 연구의 결과가 일회성이 아니라 지속적으로 도출 되도록 하는 것이다. 이를 위해 매년 혹은 분기별로 개발된 시스템에 새로운 데이터를 장착하여 지속적으로 분석함으로써 시간의 흐름에 따른 충북 관광의 이슈, 연관 관광명소, 긍/부정 분석 결과의 변화를 탐지해 내고, 그에 적합한 정책개발을 할 수 있는 체제 구축이 필요하다.

5.2. 결론

본 논문에서는 충청북도 전체 관광지를 대상으로 관련 데이터를 온라인, 오프라인 상에서 수집하고 이슈분석, 연관분석, 감성분석을 수행하였다. 특히, 지역별 (시군별, 관광지별), 계절별 (혹은 월별), 주제별로 다차원 분석과 연계하여 분석함으로써 더욱 세밀한 인사이트를 도출하고자 하였다. 이슈분석 결과 북부지역의 관광지에 대한 관심이 대부분이었으며, 중부, 남부 순서로 관심도가 높았다. 이는 남부 및 중부지역 관광지를 이슈화할 필요가 높음을 의미한다. 연관분석 결과 관광지간에 연관성과 관광지-관광키워드 (예: 숙박) 사이의 연관성을 분석하여 관광 패키지화가 가능한 근거자료를 제시하였다. 또한, 이슈분석과 연관분석에 여러 차원을 연계하여 다차원 분석을 수행한 결과 계절별로, 지역별로 어떤 이슈가 있는지, 어떤 연관관계가 있는지 확인할 수 있었고, 이를 바탕으로 더욱 세밀한 관광정책 수립이 가능하게 된다.

References

- Cho, W. S. (2014). *Demand Survey of tourism administrative and policy suggestion using Big Data*, Research Report, Chungbuk National University & Chungcheongbukdo, Korea.
- Do, H. Y., Lee, A. J. and Lee, C. W. (2013a). The study of restaurant service quality, customer satisfaction and nationality by big-data analysis. *Journal of Food Service Management Society of Korea*, **16**, 49-70.
- Do, H. Y. and Ryu, K. S. (2013b). The effect of restaurant service quality and customer satisfaction recommendation behavior by big-data analysis. *Journal of Food Service Management Society of Korea*, **16**, 233-255.
- Kong, H. S. and Song, E. J. (2013). A study on hotel CRM(Customer Relationship Management) using big data and security. *Convergence Security Journal*, **13**, 69-75.

- Lee, Y. J. and Yoon, J. H. (2014). A study on utilizing SNS big data in the tourism studies : Based on an analysis of key words for tourism information search. *International Journal of Tourism and Hospitality Research*, **28**, 5-14.
- Lee, Y. J., Jung, H. K. and Bu, S. H. (2010). Evaluation of the publicity on eTourism information : Focused on the portal site 'N'. *Journal of Tourism & Leisure Research*, **22**, 389-408.
- Moon, J. W. (2013). The utilization of ICT and implication for development of tour service: Based on case analysis of Korea and other countries. *Korea Information Society Development Institute*, **25**, 1-35.

Implementation of smart chungbuk tourism based on SNS data analysis[†]

Wan-Sup Cho¹ · Ah-Cho² · Kaaen Kwon³ · Kwan-Hee Yoo⁴

¹Department of Management Information Systems, Chungbuk National University

²³Department of Business Data Convergence, Chungbuk National University

⁴Department of Software, Chungbuk National University

Received 23 February 2015, revised 19 March 2015, accepted 24 March 2015

Abstract

With the development of mobile devices and Internet, information exchange has actively been made through SNS and Blogs. Blogs are widely used as a space where people share their experience after their visit to tourist attractions. We propose a method of recommending associated tourist attractions based on tourists' opinions using issue analysis, association analysis, and sentimental analysis for various online reviews including news in order to help to develop tour products and policies. The result shows that north area of Chungbuk province has been selected as issue attractions, and associated attractions/keywords have been identified for given well-known attraction. Positive/negative opinion for review texts has been analyzed and user can grasp the reason for the sentiments. Multidimensional analysis technique has been integrated to derive additional sophisticated insights and various policy proposal for smart tourism.

Keywords: Association analysis, big data analysis, Chungbuk tourism, issue analysis, multidimensional analysis, sentimental analysis, smart tourism.

[†] This work was supported by the research grant of the Chungbuk National University in 2012.

¹ Corresponding author: Professor, Department of MIS, Chungbuk National University, Cheongju 362-763, Korea. E-mail: wscho@chungbuk.ac.kr

² Master student, Department of Business Data Convergence, Chungbuk National University, Cheongju 362-763, Korea.

³ Master student, Department of Business Data Convergence, Chungbuk National University, Cheongju 362-763, Korea.

⁴ Processor, Department of Software, Chungbuk National University, Cheongju 362-763, Korea.