

텍스트마이닝 기법을 이용한 U-City와 Smart City의 연구 동향에 대한 분석

Study on the Trends of U-City and Smart City Researches using Text Mining Technology

임시영* · 임용민** · 이재용***
Lim, Si Yeong · Lim, Yong Min · Lee, Jae Yong

요 旨

현재 도시는 문제를 해결하고 경쟁력 향상을 위하여 ICT 기술을 접목시킨 지능형 도시에 집중하고 있다. 이러한 지능형 도시는 U-City, Smart City 등으로 정의되어 추진되고 있다. 이러한 추세 속에 U-City와 Smart City의 차이점이 무엇인지 제기되고 있다. 이에 본 연구는 타 연구와의 차별성으로 텍스트마이닝을 활용하여 논문 추세를 분석하고자 하였다. 본 연구는 지금까지 논문 추세를 통해 U-City와 Smart City의 차이점을 분석 도출하여, 향후 U-City 연구의 방향에 일조하고자 한다. 이러한 분석을 통해 국내 U-City는 도시에 대한 실제적 적용에 초점을 두고 있는 반면 Smart City는 단일 서비스 제공이나 기술개발 쪽에 초점을 두고 있다고 유추할 수 있었다. 그러나 본 연구에서는 키워드 집합의 구성에 주관적인 의견이 반영되었고, 키워드와 초록만을 분석하였다는 한계점을 지니므로 논문의 본문 또는 관련 연구보고서 등에 대한 추가적인 연구를 통해 U-City와 Smart City의 차이점을 확인하는 것이 필요하다.

핵심용어 : 지능형 도시, U-City, 유비쿼터스도시, Smart City, 텍스트마이닝, 논문추세

Abstract

City is currently developing into intelligent city which adopts the ICT technology to resolve the problems and increase the competitiveness. This intelligent city is promoted under the name of U-City or Smart City, yet it is also criticized in the trend for what the differences between U-City and Smart City are. In this study, we draws the differences between U-City and Smart City from our distinctive research method, text mining which analyzes the trend of research papers, and contribute to direction of U-City study in the future. Through this analysis, the study results in that U-City focuses practical implementation in domestic cities while Smart City focuses technological development and provision of single service. However, this paper has a limitation as the subjective opinion was reflected to configure the sets of keywords, and only keywords and abstracts were analyzed. Therefore, further studies are needed to confirm the differences between U-City and Smart City with related research papers and reports.

Keywords : Intelligent Cities, U-City, Ubiquitous City, Smart City, Text Mining, Paper Trend

1. 서 론

2010년 전 세계 도시화율은 52%에서 빠른 속도로 진전되어 2050년에는 전 세계 인구의 약 67%가 도시에서 살게 될 전망이다. 세계 도시화율은 급격한 진전

으로 나타나는 교통 혼잡, 범죄·재난, 에너지 고갈 등의 다양한 도시문제를 해결하고 도시경쟁력 향상을 위하여 도시 건설 및 관리·운영에 ICT(Information & Communication Technology; 정보통신기술)를 접목한 지능형 도시가 전 세계의 주목을 받고 있다. 이는 도시

Received: 2014.07.31, revised: 2014.08.27, accepted: 2014.09.04

* 정회원 · C&BIS, 수석컨설턴트(Member, Principal Consultant, C&BIS, tome20@naver.com)

** 국토연구원 국토정보연구본부 연구원(Assistant Research Fellow, Korea Research Institute for Human Settlements, ymlim@krihs.re.kr)

*** 교신저자 · 정회원 · 국토연구원 국토정보연구본부 책임연구원(Corresponding author, Member, Associate Research Fellow, Korea Research Institute for Human Settlements, leejy@krihs.re.kr)

인구 증가에 따라 도시 인프라를 무한대로 늘려갈 수 없기 때문에 도시 내 ICT 인프라를 적극적으로 활용하여 도시 관리의 효율성을 극대화 하는 것이 가능하기 때문이다[1].

이러한 지능형 도시는 국내에서는 U-City라는 이름으로 '08년 「유비쿼터스도시의 건설 등에 관한 법률」이 제정되면서 적극적으로 사용이 되어 왔으며 Smart City라는 용어는 '07년 애플의 아이폰 등장 이후 불기 시작한 '스마트' 열풍으로 최근 도시 전체를 대상으로 삼는 Smart City라는 용어가 국내외적으로 사용이 되고 있다[2].

지능형 도시에 대한 관심이 증대하면서 U-City와 Smart City의 차이점에 대한 논의가 점차 확대되고 있으나 이에 대한 연구는 아직까지 미흡하다. 이에 본 연구에서는 국내의 논문의 추세를 분석하여 U-City와 Smart City의 차이점을 고찰하고자 한다. 최근 빅데이터 분석이 활발히 이뤄짐에 따라 이전에는 처리할 수 없었던 많은 양의 자료에 대한 분석이 가능해졌으며 국내외의 방대한 학술적 동향을 분석하기에 적합한 기술 환경이 마련되었다.

본 연구에서는 빅데이터 기술 중 하나인 텍스트마이닝을 활용하여 지능형도시 관련 연구 동향을 분석하고자 한다. 텍스트마이닝 기법은 산업공학, 건설 등 다양한 분야에서 연구가 활발히 진행되고 있다[3]. 지능형 도시와 관련한 텍스트 마이닝 연구로는 U-City와 관련한 국내외 논문의 연구 경향을 파악하여 향후 U-City의 연구 방향을 제시한 연구가 있었고[4], 정보계측기법을 활용하여 공공데이터포털에서 수집한 공공데이터와 「제2차 유비쿼터스도시종합계획」에서 사용된 키워드를 분석·비교하여 U-City 측면에서의 공공데이터 활용성을 검토한 연구가 있다[5].

그 외 논문들의 동향을 텍스트마이닝을 활용하여 분석한 연구로는 국내 산업공학 관련 저널들에 게재된 논문들의 주제어들과 초록 내용을 바탕으로 산업공학의 주요 연구기법들의 동향을 파악한 연구가 있다[6]. 그 외 키워드 분석을 중심으로 수행된 연구는 특허를 분석하고 기술의 동향 및 차후 기술의 예측에 대한 연구[7], 한글 형태소 및 키워드 분석에 기반한 웹 문서 분류에 대한 연구[8], 관광분야의 키워드를 분석하여 숙박시설과 관광지에 대한 마케팅적 시사점을 제안한 연구[9] 등이 있다.

본 연구에서는 텍스트마이닝 기법을 활용하여 국내외 U-City와 Smart City에 관한 논문을 분석함으로써 연구동향을 파악하고자한다. 그동안 국내에서는 U-City에 대한 연구개발이 활발하게 진행되어온 반면 최근 해

외에서는 Smart City가 이를 대체하는 움직임을 보이고 있다. 이러한 상황에서 그간 U-City에 대한 연구 성과가 자칫 희석될 수가 있을 뿐 아니라, U-City라는 브랜드가치가 하락될 수 있다. 따라서 U-City와 Smart City의 연구동향을 비교함으로써 실제로 얼마나 두 개념이 차이를 보이는지, 그리고 상호 보완할 수 있는 부분이 어디인지, 또는 그간의 U-City 연구 결과를 확대하여 세계적 추세인 Smart City 연구를 선도할 수 있는 방향이 있는지에 대한 논의가 필요하다. 본 연구는 이러한 목적을 달성하기 위한 기초 연구로서 두 분야의 연구 동향을 파악하고 그 차이점을 확인하고자 한다.

2. 연구방법

본 연구에서는 'U-City'와 'Smart City'를 검색어로 설정하여 '국가과학기술전자도서관(NDSL)'에서 논문을 수집한 후 이를 분석하였다. 검색어로 'U-City'를 사용하였을 경우, 국내논문 641 편, 해외논문 94편이 검색되었으며, 'Smart City'를 사용하였을 경우, 국내논문 10편, 해외논문 237편이 검색되었다. 해당 논문의 제목을 기준으로 'U-City' 또는 'Smart City'와 관련 없는 논문들을 제외하는 작업을 거쳐, 최종적으로 'U-City'논문은 국내 591 편, 해외 70편, 'Smart City'논문은 국내 9편, 해외 227편, 총 897편을 연구대상으로 선정하였다.

연구대상이 되는 논문들을 다시 제목, 저널명, 권호 정보, 발간일, 초록, 키워드 등을 추출하였으며 이 중 발간일, 초록, 키워드를 중심으로 분석을 수행하였다. 우선 각 논문 유형별 발간추이를 살펴보고, 키워드와 초록에 대한 분석을 수행하였다. 키워드 분석은 키워드 자체를 단순 활용하기보다 유사한 유형의 키워드를 하나의 집합으로 구성한 '키워드 집합'에 대한 분석을 수행하였다. 그 결과 각 논문 유형별 키워드집합의 구성비 및 키워드집합의 연관관계를 확인할 수 있었다. 초록에 대한 분석은 초록에 수록된 단어를 우선 추출하고 일반적인 단어를 제거하는 작업을 거쳤으며 유사단어를 대표단어로 치환하여 해당 단어에 대한 연관분석을 수행하였다. 본 연구에서는 R 3.0.2를 사용하여 데이터 수집, 정리 및 분석을 수행하였다.

3. 논문 자료 분석 결과

3.1 논문 발간 추이

U-City 및 Smart City에 대한 논문의 발간 추이를 살펴보면 다음 Fig. 1과 같다. U-City 논문의 경우 국내

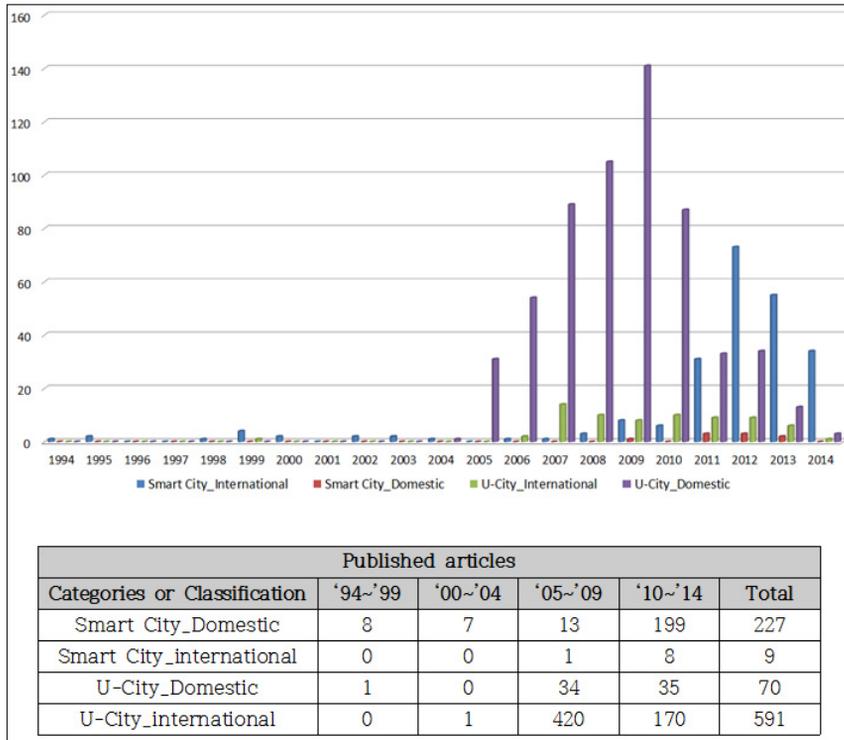


Figure 1. Changes in paper publication

논문(약 89.4%)이 주를 이루는 반면 Smart City의 경우 해외논문(약 96.2%)이 대다수를 차지하고 있음을 알 수 있다. 이는 ‘U-City’라는 용어가 해외보다는 국내에서 보편화되었으며, 역으로 해외에서는 그렇게 많이 사용하지 않고 있음을 말한다.

발간 추이를 상세히 보면, U-City 국내논문의 경우 2004년부터 나오기 시작하여 2009년에 정점을 찍었으며 이후 감소추세임을 확인할 수 있다. 이에 비해 Smart City 논문은 1994년부터 출현하였으며 초기에는 호주의 아틀레이드, 싱가포르와 같은 특정도시에 대한 내용을 다루고 있었다. Smart City 논문은 2012년에 정점을 찍고 현재는 감소하고 있는 추세임을 확인할 수 있었다. 2014년의 경우 아직 논문이 다 출간되지 않은 시점이므로, 전반적으로 추세가 감소한다기보다는 정점 이후 적정 수준을 유지한다고 볼 수도 있다.

U-City와 Smart City의 발간 추이 정점을 보면 3년 정도의 차이가 있음을 확인할 수 있으며, U-City가 국내논문에서 많이 발견되고, Smart City가 해외논문에서 많이 발견된다는 특징을 고려한다면, 만약 U-City와 Smart City가 용어의 차이만 있고 유사한 실체를 가지는 분야라면 국내의 연구가 해외보다 더욱 활발하게 선행되어 왔다고 주장할 수 있다.

3.2 키워드 분석

본 절에서는 U-City 논문과 Smart City에서 사용되는 키워드의 차이를 파악해보고자 한다. 총 논문 897편 중 저자에 의해 키워드가 제시된 논문은 391편이며, 이를 대상으로 각각 구별된 논문별로 저자가 제시한 키워드를 정리하였다. 그 결과 U-City 국내논문은 총 2,133

Table 1. The main keywords of each paper

Classification		The main keyword
U-City	Domestic paper	ubiquitous, u-service, gis, usn, u-eco city, its, spatial information, ubiquitous computing, service model, facility management, cctv
	International paper	urban informatics, south korea, seoul
Smart City	Domestic paper	u-city, Intelligent City
	International paper	internet of things, ubiquitous computing, traffic monitoring, sustainability, Smart City model, machine to machine, information and communication, governance, big data

개의 키워드(1,423 종류), U-City 해외논문은 총 46개 키워드(40종류), Smart City 국내 논문에는 총 81개의 키워드(74 종류), Smart City 해외논문은 총 224개의 키워드(196 종류)가 확인되었다. Table 1은 각 논문별 상위 키워드를 보여준다. 단, 키워드 중 동일한 내용이지만 다르게 표현된 것은 하나로 수정하여 나타내었다.

그러나 해당 키워드들은 저자가 직접 작성하기 때문에 동일한 내용의 다른 용어 사용, 한글과 영어의 혼용 등으로 그대로 분석하기에는 무리가 있다. 따라서 본 연구에서는 키워드집합을 구성하고, 전문가의 의견을 반영하여 해당 키워드들을 키워드집합에 포함시킨 후 이에 대하여 구성비 및 추이 등을 확인하고자 한다.

3.2.1 키워드 집합의 구성

U-City와 Smart City는 모두 건설을 기반으로 하여 ICT 기술을 접목시키는 분야이므로, 정보가 특정 서비스 형태로 시민에게 제공되며, 이를 위해서는 관련 시설과 기술이 필요하다. 특히 도시수준에서 이루어지기 때문에 계획요소 역시 중요한 잣대로 평가될 수 있다. 또한 U-City 및 Smart City는 기존에 있던 것이 아니므로, 이를 관련한 개념 역시 중요하므로 이를 반영하는 키워드를 따로 구분하였다. 따라서 본 연구에서는 키워드 집합으로 ‘시설’, ‘기술’, ‘서비스’, ‘계획’, ‘정의’, ‘정보’로 구분하며, 각 키워드 집합에 해당되는 키워드의 예는 Table 2와 같다. 단, 보다 유의미한 분석을 위해서 검색어로 사용된 u-City, Smart City, 이와 동일한 내용의 다른 표현들은 제외하고 총 1,656 종류의 키워드를 사용하여 키워드 집합을 구성하였다.

총 1,656 종류의 키워드 중 하나의 논문에서 동일한 키워드집합에 있는 키워드가 언급된 것을 제외하면 391편의 논문에는 총 1,024번의 키워드집합이 나타난다. 즉, 하나의 논문에 동일한 키워드집합의 키워드가 등장하는 것을 확인할 수 있으며, 각 키워드집합별로 논문에 출현하는 평균비율 파악하기 위하여 키워드집합의 출현 수정비율을 Table 3과 같이 계산하였다. 수정비율은 키워드집합의 구성비에 대해 중복성을 고려하여 표준화시킨 것이다. 예를 들어 정의라는 키워드집합은 전체 키워드의 38.6%를 차지하고 있으나, 하나의 논문에서 해당 키워드집합이 중복적으로 출현하는 비

율이 ‘54.4%(논문종류별 키워드집합 개수 합/키워드집합 내 키워드 종류)’이므로 논문 종류에 무관하게 그 비중이 동일하다고 가정하면 평균적으로 논문 중 21%에서 해당 키워드집합이 중복 사용되어 나타날 것이다. 이와 같은 방법을 각 키워드집합에 적용한 후, 표준화하면 수정비율이 33.9%로 나오고, 이는 곧, 특정 논문에서 ‘정의’라는 키워드 집합이 출현하는 비중이 평균적으로 33.9% 일 것이라는 가정이다. 평균적으로 논문 종류별 방향성에서 큰 차이가 없다면 논문 종류에 상관 없이 키워드집합은 각 논문별로 아래 표의 수정비율만큼 출현할 것이다.

Table 2. Component ratios in the keyword set

Keyword set	Examples of keyword	Composition ratio
Definition	Convergence, priority, u-GIS, citizen participation, internet of things, business model, future city	38.6% (640type)
Information	spatial information, security, Mash-up, Security, big data, digital divide, Lidar, U-City Information, traffic monitoring	15.1% (250type)
Facility	CCTV, U-City infra, U-IT Underground Facilities, 3D screen play, urban facility	8.8% (145type)
Technology, techniques	GIS, USN, ITS, Ubiquitous computing, AHP, interoperability, RFID, ontology, simulation, Wireless sensor network	16.8% (278type)
Service	U-Service, Service model, facility management, monitoring system, u-traffic, u-Zone, underground facilities management, CCTV surveillance, context-awareness service	15.8% (262type)
Plan	scenario, USP, U-City plan and design, Spatial Planning, U-eco city building guidelines	4.9% (81type)

Table 3. Revised component ratios in the keyword set

	Definition	Information	Facility	Technology, techniques	Service	Plan
Component ratios in the keyword set	38.60%	15.10%	8.80%	16.80%	15.80%	4.90%
Redundant portions in the keyword set	54.40%	58.00%	70.30%	68.00%	71.40%	65.40%
Revised component ratios in the keyword set	33.9%	14.2%	10.0%	18.5%	18.2%	5.2%

3.2.2 키워드 집합별 구성비

Table 4는 각 논문별 키워드집합 출현 여부의 비율을 보여주고 있다. 예를 들면 Smart City의 국내논문(S_D) 중 약 32.0%에는 ‘정의’를 나타내는 키워드가 등장한다. 이를 위 Table 4와 비교해보면 논문 종류별로 나타나는 키워드 집합의 출현 비율을 확인할 수 있다.

이를 그림으로 표현하면 다음 Fig. 2와 같다. 해당 그림은 키워드 집합의 구성 시 반영된 키워드집합 출현 수정비율을 각 키워드 집합의 기준으로 정하고, 상대적 크기를 나타낸 그림이다. U-City 국내논문의 경우는 각 키워드 집합의 비율이 기준에 근접하고 있으나, U-City 해외 논문은 정의 및 시설 키워드집합의 비중이 높고, 정보 및 서비스, 기술의 키워드 집합 비율이 낮다.

Table 4. Appearance rates of the keyword set

	S_D	S_I	U_D	U_I
Definition	32.0%	34.0%	33.8%	47.4%
Information	20.0%	13.6%	14.3%	5.3%
Facility	12.0%	9.7%	9.7%	21.1%
Technology, techniques	24.0%	13.6%	19.0%	10.5%
Service	12.0%	25.2%	17.8%	10.5%
Plan	0.0%	3.9%	5.5%	5.3%

Smart City 국내 논문의 경우, 정보, 기술 및 기법은 기준 이상의 비율을 가지나, 정의, 계획, 서비스 집합은 기준 보다 낮은 비율로 구성되어 있으며, Smart City 해외 논문은 서비스가 기준 이상의 비율을 차지하고 있으며, 기술기법에 대한 비율이 상대적으로 낮게 나타났다.

3.2.3 논문별 키워드 집합의 연관규칙 검토

연관분석은 장비구니분석으로도 불리는 데이터마이닝 기법의 하나이다. 본 연구에서는 키워드집합에 대한 연관분석을 수행하여 논문 종류별 각 키워드집합간의 연관규칙을 확인해보고자 한다. 연관분석의 척도로는 지지도(support), 신뢰도(confidence), 향상도(lift)라는 것이 사용된다.

지지도는 전체 문서에서 키워드 A와 B가 함께 나오는 확률을 의미하고($P(A \cap B)$), 신뢰도는 키워드 A가 나오는 관계 속에서 A와 B가 동시에 나온 비율을 의미하며($P(B|A)$), 향상도는 두 키워드가 독립인지 여부를 판단하는 척도이다($P(A \cap B)/(P(A) \cdot P(B))$). 향상도의 경우 1보다 클수록 연관성이 크다고 판단한다. 즉, 향상도가 클수록 키워드 X가 나오에 따라 키워드 Y가 나올 확률이 커진다는 것을 의미한다. 이 두 키워드의 관계는 ‘X->Y’라고 표현할 수 있는데 그 의미는 하나의 논문에서 X라는 키워드가 나왔을 때 Y라는 키

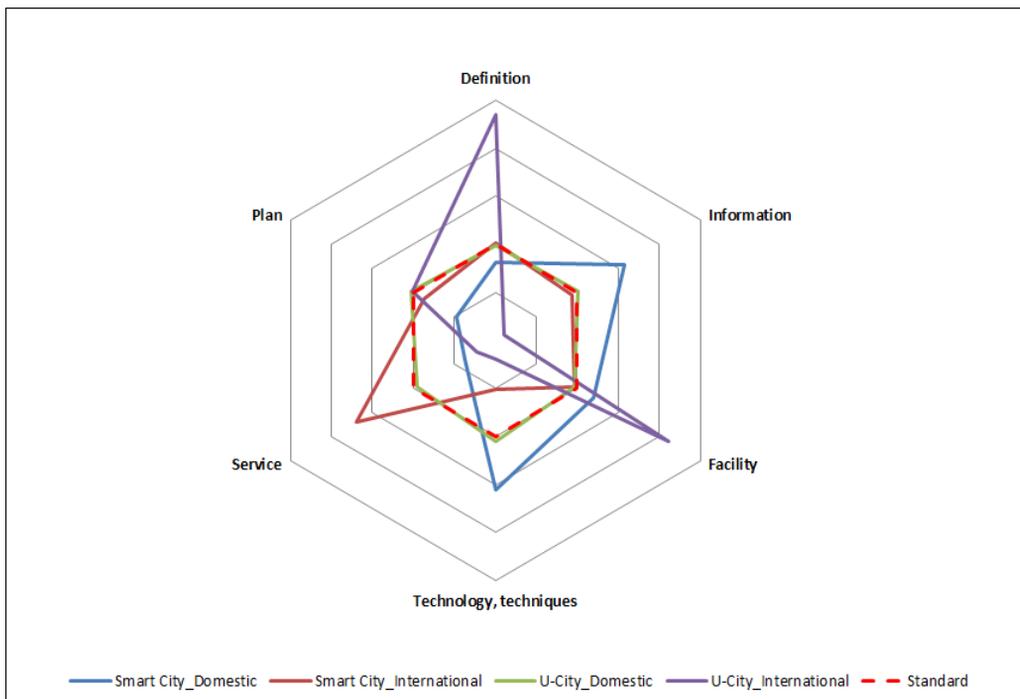


Figure 2. Component ratio of keyword set by type of paper

워드스가 나온다는 연관관계를 표시하는 것이다.

본 연구에서는 적절한 수의 규칙이 도출될 수 있도록 지지도 0.1 이상, 신뢰도 0.3 이상의 연관규칙을 조사하여 이 중 원소가 하나씩인 규칙을 그래프화 하였다. 아래 Figure3, 4는 이를 나타낸 것으로 그래프에서 선이 두꺼워질수록 두 키워드집합의 관계의 지지도 값이 커지며, 색이 진해질수록 향상도 값이 커진다.

Smart City 국내 논문의 경우, ‘정의<->기술 및 기법’ 간의 지지도 값이 0.56으로 가장 컸으며, 그때의 향상도 값은 0.94로 나타났다. 또한 ‘서비스<->정보’, ‘시설<->정보’간은 지지도 값이 0.22였으나 향상도 값이 1.2로 가장 크게 나타났다. 즉, 정의와 기술 및 기법이 하나의 논문에 가장 많이 동시 출현했으며, 서비스 키워드집합의 키워드가 나타나면 정보 키워드집합의 키워드(또는 그 반대)가 나타나는 확률이 높으며, 시설 키워드집합의 키워드가 나타나면 정보 키워드집합의 키워드(또는 그 반대)가 나타날 확률이 높았다.

Smart City 해외논문의 경우, 마찬가지로 ‘정의<->기술 및 기법’ 간의 지지도 값이 0.59로 가장 컸으며, 그때의 향상도 값은 0.99로 나타났다. 또한 ‘서비스<->정보’는 지지도 값이 0.18이었으나 향상도 값이 1.39로 가장 크고, ‘계획->정의’와 ‘시설 ->정의’가 지지도 0.26, 향상도 1.11로 그 다음으로 큰 향상도 값을 보인다. 즉, 정의와 기술 및 기법이 하나의 논문에서 동시에 출현한 경우가 가장 많았으며, 서비스 키워드집합의 키워드가 나타나면 정보 키워드집합의 키워드(또는 그 반대)가 나타나는 확률이 높고, 계획 또는 시설 키워드집합의 키워드가 나타나면 정의 키워드집합의 키워드(또는 그 반대)가 나타날 확률이 높았다.

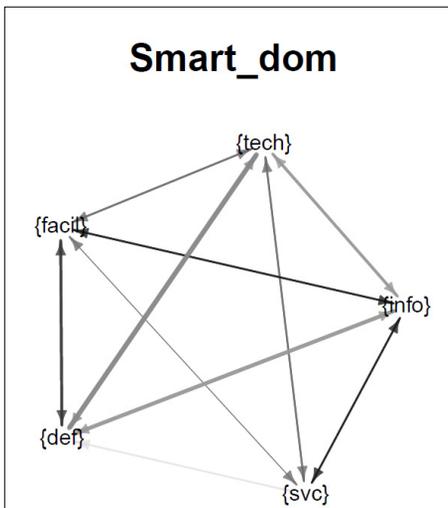


Figure 3. Smart City Domestic

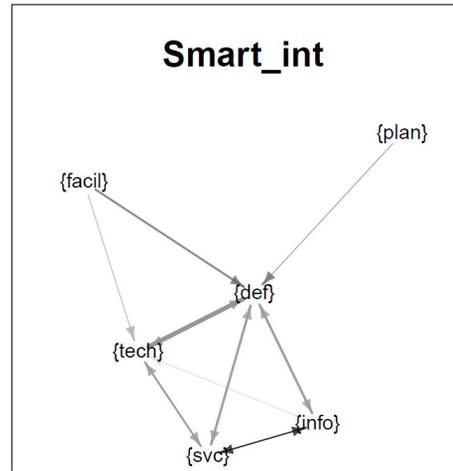


Figure 4. Smart City International

U-City 국내 논문의 경우, ‘정의<->기술 및 기법’, ‘정의<->서비스’ 간의 지지도 값이 0.41로 가장 컸으며, 그때의 향상도 값은 0.93으로 나타났다. 또한 ‘시설->서비스’의 지지도는 0.14였으나 향상도 값이 1.13으로 가장 높게 나타났다. 이는 곧, U-City 국내논문의 경우, 정의와 기술 및 기법, 정의와 서비스가 하나의 논문에 가장 많이 동시 출현했으며, 시설 키워드집합의 키워드가 나타나면 서비스의 키워드집합의 키워드가 나타날 확률이 높았음을 의미한다.

U-City 해외논문의 경우, ‘정의<->기술 및 기법’ 간의 지지도 값이 0.59로 가장 컸으며, 그때의 향상도 값은 0.99로 나타났다. 또한 ‘서비스<->정보’는 지지도 값이 0.18이었으나 향상도 값이 1.39로 가장 크고, ‘시설->정의’가 지지도 0.44로 가장 크고, 그때의 향상도

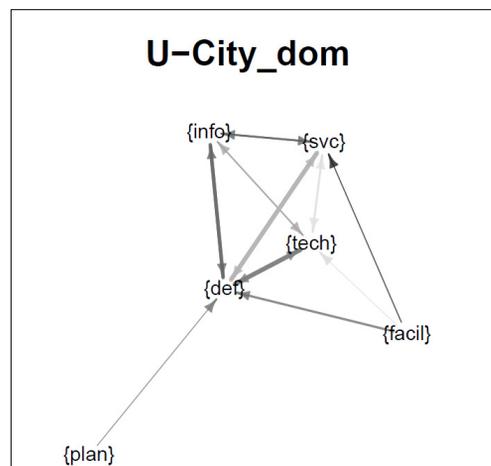


Figure 5. U-City Domestic

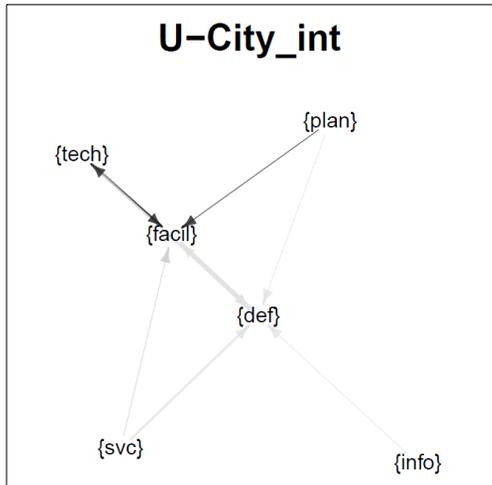


Figure 6. U-City International

값은 1.0으로 나타났다. 향상도의 관점에서 보면, ‘시설 <-> 기술 및 기법’과 ‘계획>시설’이 2.25로 가장 높게 나타났다. 즉, U-City 해외논문의 경우, 정의와 기술 및 기법 키워드집합의 키워드가 동시에 출현한 경우가 가장 많았으며, 시설 키워드집합의 키워드가 나타나면 기술 및 기법 키워드집합의 키워드(또는 그 반대)가 나타나는 확률이 높았으며, 계획 키워드집합의 키워드가 출현한 논문에는 시설 키워드집합의 키워드가 출현하는 확률이 높았다.

3.3 초록 분석

본 절에서는 논문에서 제시된 초록(abstract)에 출현하는 단어들에 대한 연관규칙을 분석하고자 한다. U-City 논문의 경우 국내논문 293편, 해외논문 70편, Smart City 논문의 경우, 국내논문 9개 해외논문 154편을 대상으로 하였으며 국내논문의 경우, 한글 요약만 있는 것은 제외하였다. 키워드 집합의 연관분석과 동일한 방법을 사용하였으며, 적절한 수의 연관규칙을 확인하기 위하여 이때 기준으로 Smart City 국내논문의 경우 지지도/신뢰도를 각각 0.3으로 설정하였고, 나머지는 지지도/신뢰도를 각각 0.2로 설정하였다. 그 결과 U-City 국내논문의 경우 총 105개의 연관규칙이 도출되었으며 이 중 원소가 1개씩인 연관규칙 58개를 정리하였으며, U-City 해외논문의 경우 총 119개 중 원소가 1개씩인 연관규칙 56개, Smart City 국내논문의 경우 총 565개의 연관규칙 중 원소가 1개인 연관규칙 94개, Smart City 해외논문의 경우 총 170개의 연관규칙 중 원소가 1개인 연관규칙 84개를 정리하였다. 그 결과는 아래 Fig. 7, 8, 9, 10과 같다. 그림에서 LHS는 선행하

는 단어이며, RHS는 그에 따라 후행하는 단어를 의미한다. 원의 크기가 크면 지지도 값이 크고, 원의 색깔이 짙을수록 향상도 값이 크다. Fig. 7과 8을 보면, U-City 국내논문의 경우 U-City라는 단어가 타 단어와 연관성이 크기는 하지만, 다른 단어들 사이에도 연관성을 갖는데 비해, Smart City의 경우, 대부분의 단어가 ‘Smart’와 ‘City’에만 연관성을 갖는 것을 알 수 있다. 이점은 U-City 논문에 비해 Smart City 논문이 보다 세분화되어 있다고 볼 수 있으며, 또는 Smart City 논문이 U-City에 비해 내부 구성요소간의 관계를 고려하는 단계까지 발전하지 못했다고도 볼 수 있다.

U-City 국내논문의 경우, ‘U-City <-> service’가 0.45로 가장 높은 지지도 값을 보이며 그때의 향상도 값은 1.042이다. 향상도 기준으로 보면 ‘urban <-> city’가 1.316으로 가장 높고(지지도는 0.201), 다음으로 ‘service <-> technology’가 1.198(지지도는 0.307), ‘information <-> system’이 1.192(지지도는 0.215)로 나타났다.

U-City 해외논문의 경우, ‘u-city <-> information’이 0.3755로 가장 높은 지지도 값을 보이며 그때의 향상도 값은 1.167이다. 향상도 기준으로 보면 ‘system <-> information’이 1.684로 가장 높고(지지도는 0.214), 다음으로 0.214), ‘propose <-> ubiquitous’가 1.445(지지도는 0.286), ‘ubiquitous <-> computing’이 1.445(지지도는 0.214), ‘technology <-> city’가 1.441(지지도는 0.25)로 나타났다.

Smart City 국내논문의 경우, ‘u-city <-> service’가 0.667로 가장 높은 지지도 값을 보이며 그때의 향상도 값은 1.000이다. 향상도 기준으로 보면 ‘integrate <-> ubiquitous’가 3.000으로 가장 높고(지지도는 0.333), 다음으로 ‘environment <-> information’이 2.250(지지도는 0.444), ‘management <-> urban’, ‘future <-> concept’ 등이 2.25(지지도는 0.333)로 나타났다.

Smart City 해외논문의 경우, ‘smart <-> city’가 0.87로 가장 높은 지지도 값을 보이고, 그때의 향상도 값은 1.061로 나타났다. 향상도 기준으로 보면, ‘management <-> system’이 1.256(지지도는 0.292)가 가장 높게 나타났고, 다음으로 ‘information <-> smart’가 1.124(지지도는 0.364), ‘technology <-> smart’가 1.124(지지도는 0.247) 순으로 나타났다.

U-City 국내논문에서 가장 높은 지지도 값을 가진 ‘u-city <-> service’와 대응하는 Smart City 해외논문의 ‘smart <-> service’는 지지도 0.396, 향상도 1.032 값을 나타내고 있다. 역으로 Smart City 해외논문에서 가장 높은 지지도 값을 가진 ‘smart <-> city’는 U-City

국내논문에서는 ‘u-city <-> city’로 대응될 수 있으며 지지도 0.416, 향상도 1.029로 나타났다.

U-City 국내논문에서 가장 높은 향상도 값을 가진 ‘urban <-> city’는 Smart City 해외논문에서 지지도

0.253, 향상도 1.007 값을 나타내고 있다. 역으로 Smart City 해외논문에서 가장 높은 향상도 값을 가진 ‘management <-> system’은 U-City 국내논문에서는 지지도/신뢰도 기준을 통과하지 못하였다.

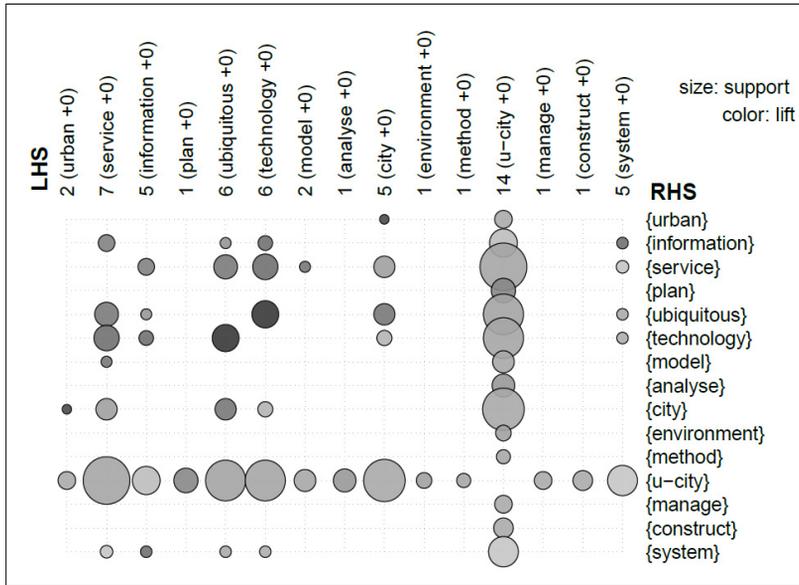


Figure 7. Association rules in U-City domestic papers

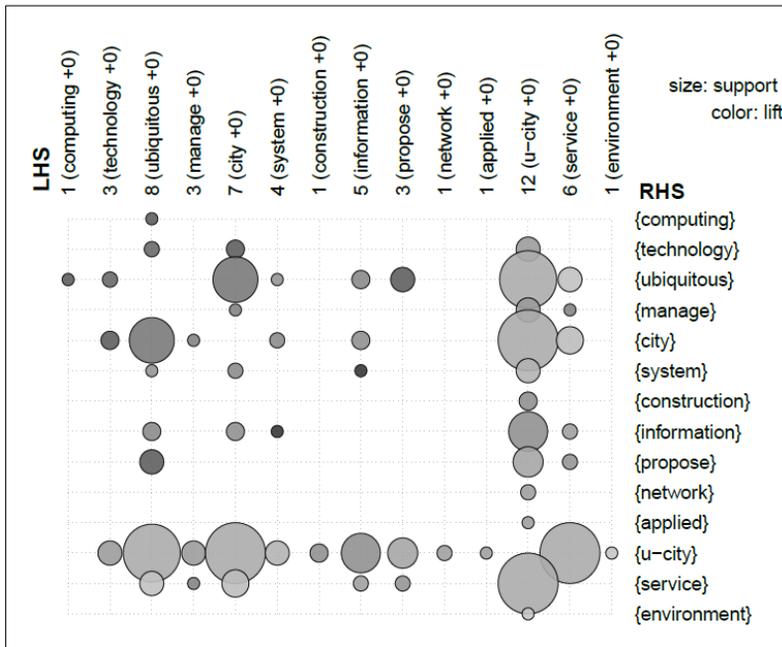


Figure 8. Association rules in U-City international papers

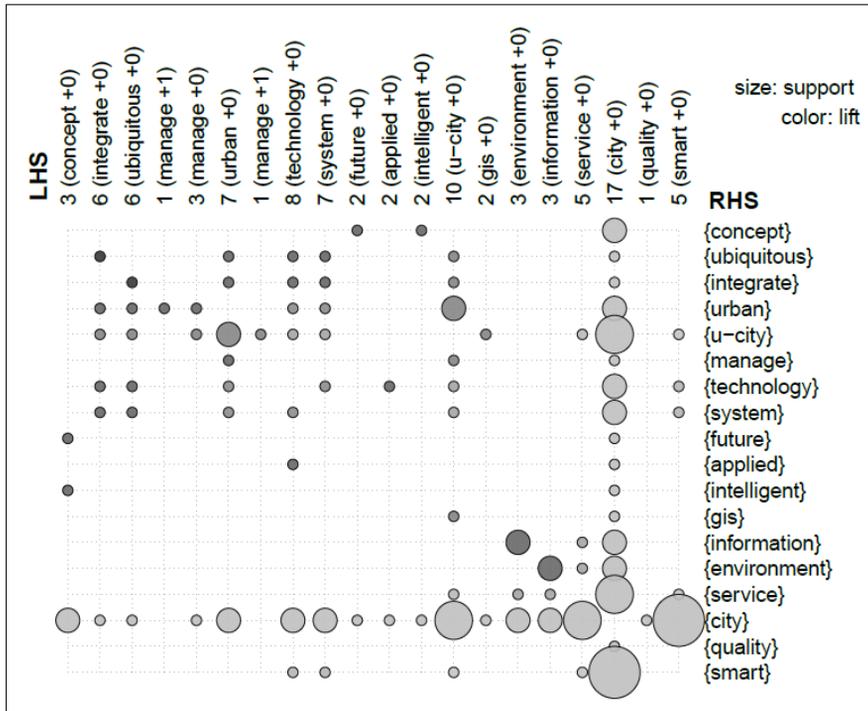


Figure 9. Association rules in Smart City domestic papers

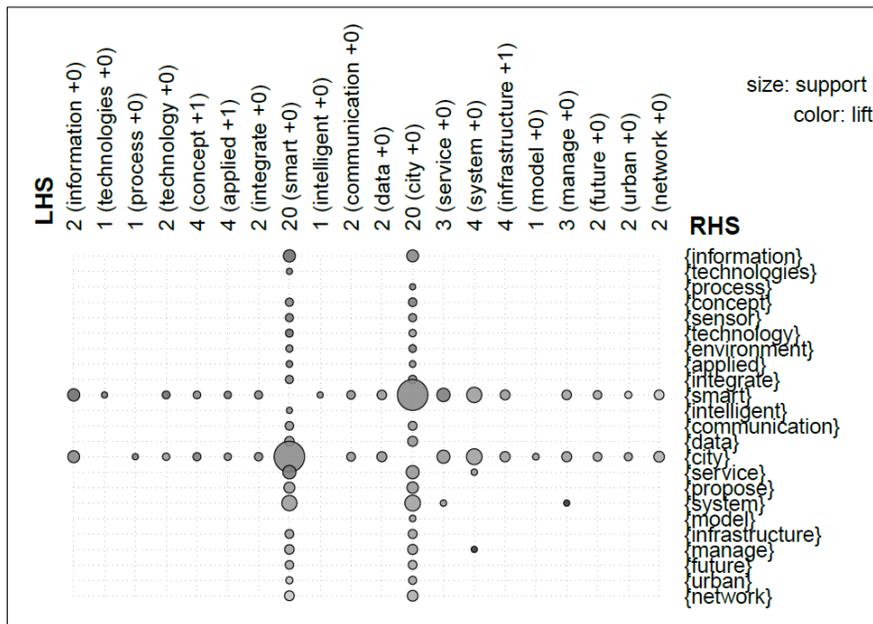


Figure 10. Association rules in Smart City international papers

Table 5. Comparison of the associated keyword between U-City & Smart City

Division	U-City				Smart City			
	RHS	support	confidence	lift	RHS	support	confidence	lift
Common	model	0.283	0.349	1.032	model	0.201	0.226	0.968
	technology	0.406	0.500	1.039	technology	0.247	0.277	1.124
	environment	0.242	0.298	1.053	environment	0.227	0.255	1.063
	urban	0.256	0.315	1.026	urban	0.227	0.255	0.937
	management	0.256	0.315	1.026	management	0.292	0.328	1.012
	information	0.324	0.399	0.975	information	0.364	0.409	1.124
	service	0.451	0.555	1.042	service	0.396	0.445	1.088
	system	0.338	0.416	0.945	system	0.448	0.504	1.021
	city	0.416	0.513	1.029	city	0.870	0.978	1.061
Distinction					method	0.232	0.286	1.034
					analyse	0.290	0.357	1.079
					construction	0.270	0.332	1.024
					plan	0.300	0.370	1.129
					ubiquitous	0.406	0.500	1.046
					concept	0.253	0.285	1.069
					application	0.214	0.241	1.091
					intelligent	0.201	0.226	1.056
					sensor	0.253	0.285	1.096
					integrate	0.253	0.285	1.069
				communication	0.266	0.299	1.047	
				future	0.266	0.299	1.002	
				infrastructure	0.273	0.307	1.026	
				data	0.292	0.328	1.032	
				network	0.299	0.336	0.940	

Table 5는 ‘U-City’와 ‘Smart’와 연관관계를 가지는 단어들을 나열한 것이다. 두 단어와 공통적으로 연관이 있는 단어들은 ‘model, technology, environment, urban, management, information service, system, city’이다. 이에 비해 U-City와 연관이 있으나 Smart와 연관이 없는 단어는 method, analyse, construction, plan, ubiquitous가 있으며 Smart와 연관이 있으나 U-City와 연관이 없는 단어는 ‘concept, application, intelligent, sensor, integrate, communication, network’등으로 나타났다(보다 정확히 표현하면 연관관계가 없는 것이 아니라 연관관계가 지지도/신뢰도 기준을 넘지 못한 것임). 이점을 고려하면, U-City 논문이 방법론, 분석, 건설, 계획 등에 초점을 맞추고 있는 반면, 상대적으로 Smart City 논문은 개념, 응용, 지능화, 기술 등에 초점을 맞추고 있다고 볼 수 있다.

U-City에서 information과 system, Smart City에서 network의 향상도 값이 1보다 작으므로 음의 연관관계를 가지며 나머지는 모두 양의 연관관계를 갖는 것을 알 수 있다. 공통부분에서 model, urban, management를 제외하면 모든 부분에서 Smart City의 향상도 값이 높다. 이 점은 Smart City 논문에서 해당 단어들에 대해 보다 더 예측력이 있다고 볼 수 있다 (Smart City라는 단어가 나왔으면, 해당 논문에는 technology가 나올 가능성이 U-City 논문보다 크다는

의미이다).

4. 결론

본 연구에서는 “U-City”와 “Smart City”를 검색어로 하여 국내의 관련 논문을 수집하고 해당 논문들의 키워드 및 초록에 대한 분석을 수행하였다. 이를 통해 U-City와 Smart City에 대한 차이점을 확인하고자 하였다.

우선 논문 비중 및 발간주기를 보면 U-City의 경우 국내논문이 많고, 2005년부터 시작되어 2009년을 전성기로 보였으며, Smart City의 경우 해외논문이 다수를 차지하고 있으며 1994년에 출현하였다가 2000년대 후반부터 증가하기 시작하여 2012년을 최고점을 통과하였다.

키워드 분석을 위하여 6개의 키워드집합을 구성하고 해당 키워드 집합의 차이점을 비교하였으며, 그 결과 U-City 국내 논문에 비하여 Smart City 해외 논문은 서비스와 관련된 키워드집합의 비중이 컸으며, 기술 및 기법, 계획 키워드집합의 비중이 평균적으로 작았다.

키워드집합의 연관규칙을 살펴보면 U-City 국내 논문의 경우, ‘시설’ 키워드집합의 키워드가 나타나면 ‘서비스’ 키워드집합의 키워드가 나타날 가능성이 높

았으며, Smart City 해외 논문의 경우, ‘서비스’ 키워드 집합의 키워드와 ‘정보’ 키워드 집합의 키워드가 상호 예측력이 가장 높았다.

초록에 대한 연관분석 결과를 살펴보면, U-City 국내 논문의 경우, 다양한 단어들의 연관성이 나타난 반면, Smart City 해외 논문의 경우 타 단어들이 ‘Smart’와 ‘City’에 깊은 연관성을 나타내고 있다.

또한 U-City 국내 논문과 Smart City 해외 논문에 대하여 공통 또는 차이가 나는 단어에 대한 연관성을 살펴보면, 모델, 기술, 서비스, 운영, 정보 등에서는 공통적으로 연관성을 보이고 있으나 U-City 국내 논문의 경우 분석, 건설, 계획 등의 단어가 등장하는 반면 Smart City 해외 논문의 경우 개념, 응용, 기술적 단어들이 등장하고 있다.

이러한 점은 국내 U-City는 도시에 대한 실제적 적용에 초점을 두고 있는 반면 Smart City는 단일 서비스 제공이나 기술개발 쪽에 초점을 두고 있다고 유추할 수 있다. 그러나 본 연구에서는 키워드 집합의 구성에 주관적인 의견이 반영되었고, 키워드와 초록만을 분석하였다는 한계점을 지니므로 논문의 본문 또는 관련 연구보고서 등에 대한 추가적인 연구를 통해 U-City와 Smart City의 차이점을 확인하는 것이 필요하다.

감사의 글

이 논문은 국토교통과학기술진흥원 도시건축연구사업(과제번호: 13AUDP-B070713-01)의 지원을 받아 작성되었습니다.

References

1. National IT Industry Promotion Agency, 2013, Smart city building national and international trends and implications, separate volume,
2. National IT Industry Promotion Agency, 2013, Future city of the fusion of cutting-edge industries and ICT - smart city, New growth engines industry and fusion web magazine.
3. Bae, K. Y, Park, J. H, Kim, J. S, Lee, Y.S, 2013, Analysis of the abstracts of research articles in food related to climate change using a text-mining algorithm, Journal of the Korean Data & Information Science Society, 24(6), pp.1429-1437.
4. Kim, K, Chun, J. Y, Shin, D. B, Lim, S. Y, 2011, Research trends of U-City theses in Korea and oversea, Journal of the Korean Society for Geospatial Information System, 19(1), pp.53-61.
5. Jeong, D. W, Min, K.J, Yi, M. S, Shin, D.B, 2014, A study on the possibility of the use of Public-Data in U-City, 2014 Korean Society for Geospatial Information System conference, pp.179-182.
6. Cho, G. H, Lim, S. Y, Hur, S, 2014, An analysis of the research methodologies and techniques in the industrial engineering using text mining, Journal of the Koran Institute of Industrial Engineers, 40(1), pp.52-59
7. Choi, J. H, Kim, H. S, Lim, N. G, 2011, Keyword network analysis for technology forecasting, Journal of the Korea Intelligent Information System Society, 17(4), pp.227-240.
8. Park, D. H, Choi, W. S, Kim, H. J, Lee, S. L, 2012, Web document classification based on hangeul morpheme and keyword analyses, Journal of the Korea Information Processing Society, 19(4), pp.263-270.
9. Lee, Y. J, 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”, Journal of the Korea Tourism Research Association, 28(3), pp.5-14.