

이벤트와 관련된 주변 관광지 자동 추천 알고리즘 개발

안진현¹, 임동혁^{2*}

¹제주대학교 경영정보학과, ²호서대학교 컴퓨터공학과

Automatic Recommendation of Nearby Tourist Attractions related to Events

Jinhyun Ahn¹, Dong-Hyuk Im^{2*}

¹Department of Management Information Systems, Jeju National University

²Department of Computer Engineering, Hoseo University

요약 관광객이 관광 도중에 각종 문화제, 전시회, 공연 등의 이벤트에 참여하는 경우가 있다. 관광객이 이벤트에 참여 후 다음 관광지를 결정하게 되는데, 관광지 정보를 얻을 수 있는 수단으로 지도 서비스, 블로그와 같은 소셜네트워크서비스 등이 존재한다. 지도 서비스를 활용하면 관광객이 현재 위치한 장소 주변의 관광지를 쉽게 검색할 수 있다. 이는 위치 기반 관광지 추천으로 활용될 수 있다. 블로그 등은 관광지의 내용을 담고 있기 때문에 관광객이 이벤트의 내용과 관련된 관광지를 찾을 수 있다. 이는 내용 기반 관광지 추천으로 활용될 수 있다. 하지만, 위치 기반 추천의 경우 이벤트의 내용과 관련이 없이 단순히 가까운 관광지가 추천이 될 수 있고, 내용 기반 추천의 경우 거리가 먼 관광지가 추천이 될 수 있는 단점이 있다. 위치와 내용을 모두 고려하는 관광지 추천 서비스는 거의 없다. 본 연구에서는 두 가지 방법의 장점을 취하기 위해 한국관광공사 LOD(Linked Open Data), 위키피디아, 국어사전 등에 기반하여 위치와 내용을 모두 고려한 관광지 추천 알고리즘을 제시한다. 관광지의 설명글로부터 명사들을 추출한 뒤 다른 관광지의 명사들과 비교를 하여 동일한 명사가 많이 있을수록 내용이 관련이 있다고 판단한다. 정확히 동일한 명사가 없어도 위키피디아에 있는 키워드를 활용하여 관련된 명사가 존재할 경우에도 관련이 있다고 판단한다. 각 관광지의 위도와 경도를 기준으로 거리를 계산한 뒤 사용자가 선택한 가중치로 상기 내용 기반 관련도와 선행결합하여 추천순위를 계산한다.

Abstract Participating in exhibitions is one of the major activities for tourists. When selecting their next travel destination after participating in an event, they use map services and social network services, such as blogs, to obtain information about tourist attractions. The map services are location-based recommendations, because they can easily retrieve information regarding nearby places. Blogs contain informative content about tourist attractions, thereby providing content-based recommendations. However, few services consider both location and content. In location-based recommendations, tourist attractions that are not related to the content of the event attended might be recommended. Content-based recommendation has a disadvantage in that events located at a distance might get recommended. We propose an algorithm that considers both location and content, based on information from the Korea Tourism Organization's Linked Open Data (LOD), Wikipedia, and a Korean dictionary. By extracting nouns from the description of a tourist attraction and then comparing them with nouns about other attractions, a content-based relationship is determined. The distance to the event is calculated based on the latitude and longitude of each tourist attraction. A weight selected by the user is used for linear combination with the content-based relationship to determine the preference order of the recommendations.

Keywords : Tourist Attractions Recommendation, Location-based, Content-based, Ontology, RDF

이 논문은 2018년도 제주대학교 교원성과지원사업에 의하여 연구되었음

*Corresponding Author : Dong-Hyuk Im(Hoseo University)

email: dhim@hoseo.edu

Received November 27, 2019

Revised December 31, 2019

Accepted March 6, 2020

Published March 31, 2020

1. 서론

관광을 하나 이상의 장소(관광지)를 방문하여 풍경, 문물 등을 구경하는 행위라고 보면, 어떤 관광지에서 다음 관광지를 적절하게 선택하는 것이 중요하다. 관광객 개인이 여러 관광지를 비교 분석하여 가장 관심이 있는 관광지를 선택하는 방법이 있을 수 있는데 관광객 개인의 배경지식의 부족으로 원하는 관광지를 찾지 못할 수도 있다. 관광 전문가로부터 관광지를 추천받는 방법도 있지만 마찬가지로 관광 전문가의 배경지식에 의해 제한이 될 뿐만 아니라 관광 도중에 관광 전문가에게 지속적으로 추천을 받기는 어렵다. 이러한 문제를 해결하기 위해서는 스마트폰 등의 이동형 단말기를 통해 언제 어디서나 관광지를 추천받을 수 있는 시스템이 필요하다.

기존 관광지 추천 시스템은 크게 위치 기반 관광지 추천과 내용 기반 관광지 추천으로 나뉜다[1]. 위치 기반 관광지 추천은 사용자가 현재 위치한 주변의 관광지를 추천하는 것으로 위도/경도로 거리를 계산하여 관광지들의 목록을 거리가 짧은 순으로 제시하는 방법이다. 대표적인 사례로 카카오맵, 네이버맵 등의 지도 서비스이다. 현재 위치한 장소 주변의 관광지들이 지도에 시각화 되고 거리도 확인할 수 있다. 또한, 문화체육관광부에서 운영하는 문화포털(<https://www.culture.go.kr>)의 경우에도 각종 행사 정보를 제공하면서 주변 관광지에 대한 설명도 제공한다. 내용 기반 관광지 추천은 관광지의 주제, 역사 등이 연관된 다른 관광지를 추천하는 방법이다. 관광지가 주제 또는 역사를 기준으로 분류가 돼 있는 경우 바로 유사한 관광지를 찾을 수 있지만 공식화된 분류 기준이 있는 것이 아니기 때문에 관광지의 설명글을 분석하는 방법을 사용할 수 있다[3]. 내용 기반 관광지 추천은 아직 연구단계로 서비스화된 사례가 거의 없지만, 기존 지도 서비스에 쉽게 통합될 수 있기 때문에 향후 관련 서비스가 개발될 것으로 기대된다.

위치 기반 관광지 추천의 경우 현재 위치한 관광지의 근처에 있는 관광지를 추천하기 때문에 관광객이 바로 이동해서 추천된 관광지를 바로 방문할 수 있는 장점이 있는 반면 현재 위치한 관광지와 전혀 관련이 없는 경우 관광객의 흥미를 유발하기 어려운 단점이 있다. 내용 기반 관광지 추천의 경우 현재 위치한 관광지의 주제 또는 역사와 연관된 관광지를 추천하기 때문에 관광객의 흥미를 끌 수 있지만 이동 거리를 고려하지 않기 때문에 이동의 측면에서 비현실적인 추천이 될 수 있는 단점이 있다.

본 연구에서는 이러한 위치 기반 관광지 추천과 내용

기반 관광지 추천의 장점만을 취하기 위해 이동 거리와 내용 유사도를 모두 고려하는 관광지 추천 방법을 제안한다.

2. 관련연구

기존 관광지 추천에 관련된 연구는 설문조사 기반 추천, 위치 기반 추천, 그리고 내용 기반 추천으로 나눌 수 있다. 설문조사 기반 추천은 설문조사를 바탕으로 개인 특징별 선호 관광지 목록을 도출하는 방법이다[2,6,7,8,9]. 설문조사라는 제한된 데이터에 기반하기 때문에 추천의 다양성에 한계가 있다. 위치 기반 추천은 위도/경도 정보를 바탕으로 인접해 있는 관광지를 추천하는 방법으로 차 네비게이션, 한국관광공사 웹서비스 등에서 제공한다. 한국관광공사의 경우 선택한 관광지의 “주변 가볼만한 곳” 메뉴에서 주변 관광지 목록을 제공한다[16]. 선택한 관광지와 전혀 관련이 없는 관광지도 연관된 관광지로 추천이 될 수 있다는 문제점이 있다. 내용 기반 추천은 선택한 관광지가 담고 있는 내용과 연관된 관광지를 추천하는 방법이다[3]. 선택한 관광지와 이야기를 통해 연결된 관광지를 추천할 수 있지만 거리가 먼 관광지를 추천할 수 있다는 단점이 있다.

또한, 내용 기반 추천의 일종인 관광객에 대한 정보를 기반으로 관광지를 추천하는 연구도 활발하다[16]. 관광객과 관광지간의 유사도를 기반으로 추천하는 연구에서는 관광객이 과거에 경험했던 관광지와 관련된 키워드와 관광지의 키워드를 코사인 유사도(Cosine Similarity)를 기반으로 추천을 한다[17]. 취향이 유사한 관광객 그룹을 구성하고 이 관광객 그룹과 관광지간의 유사도를 측정하여 추천하는 연구도 있지만 취향이 유사한 관광객 데이터가 충분히 쌓여 있어야만 가능하다는 단점이 있다[18].

국내에서 접근할 수 있는 관광 관련 서비스는 다음과 같다. 한국문화관광연구원에서 운영하는 관광지식정보시스템(www.tour.go.kr)에서는 대한민국의 관광 관련 통계를 오픈 API와 파일 형태로 제공한다. 출입국 통계, 여행 실태조사 등 정부 차원에서 생산할 수 있는 데이터를 얻을 수 있다. 제주데이터허브(www.jejudatahub.net)는 제주 관광 관련 민관 데이터를 통합하여 제공한다. 한국정보화진흥원이 추진한 2017년 빅데이터 플래그십 시범사업을 통해 구축됐으며 현재는 카카오에서 운영하고 있다. 2018년 5월 현재 655개의 데이터셋을 제공하고 있으며, 검색 키워드 순위, 자동차 이동경로, 관광객 방문

통계 등의 정보를 제공한다. 제주 스마트 관광 플랫폼(jstp.jejutour.go.kr)은 비콘(Beacon)이 설치된 제주도 내 주요 관광지에서 수집된 정보를 제공하는 서비스이다. 관광객의 방문패턴과 이동경로 데이터를 오픈 API나 웹 브라우저를 통해 실시간으로 얻을 수 있다.

3. 관광지 자동 추천

본 연구는 관광지 관련 데이터에 기반하여 관광지를 자동으로 추천하는 방법에 관한 것이다. 관광지 관련 데이터는 한국관광공사에서 관리하고 있는 한국관광공사 LOD(Linked Open Data)를 활용한다. 한국관광공사 LOD는 관광분야의 지식을 표현한 온톨로지의 일종으로 온톨로지는 차세대 웹 환경의 핵심 표현 방법들 중 하나이다[12,14].

온톨로지는 지식을 표현하는 방법론인데 기술적인 의미에서 온톨로지는 RDF(Resource Description Framework) 트리플의 집합이다. RDF 트리플은 주어(Subject), 술어(Predicate), 목적어(Object)로 구성된다. 주어는 개체에 대한 식별자이고 개체의 속성은 술어로 기술하고 속성의 값은 목적어 부분에 표현한다. 개체는 속성과 속성의 값의 쌍의 집합으로 표현할 수 있기 때문에 여러 개의 RDF 트리플로 하나의 개체에 대한 지식을 표현할 수 있다. 예를 들어, 한라산은 제주도에 위치해 있고 제주도의 넓이는 1849km^2 이다 라는 지식은 다음과 같이 마침표(.)로 구분된 2개의 RDF 트리플로 표현이 된다.

```
<한라산> <위치> <제주도> .
<제주도> <넓이> "1849km2" .
```

첫 번째 RDF 트리플에서 <제주도>가 목적어 부분에 있는데 두 번째 RDF 트리플의 주어이기도 하다. 제주도 넓이의 경우 숫자로 표현이 되는데 목적어 부분에 위치한다. 이와 같이 주어와 술어에는 개체 또는 술어의 식별자인 URI(Uniform Resource Identifier)가 가능하고 목적어에는 URI 또는 문자열, 숫자열 등의 리터럴(Literal)이 가능하다.

```
PREFIX                                rdf:
<http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX kto: <http://data.visitkorea.or.kr/ontology/>
PREFIX ktop: <http://data.visitkorea.or.kr/property/>
PREFIX                                wgs:
<http://www.w3.org/2003/01/geo/wgs84_pos#>

SELECT *
WHERE
{
  ?s rdf:type kto:Attraction .
  ?s rdfs:label ?s_label .
  ?s dc:description ?desc .
  ?s ktop:category ?category .
  ?s wgs:lat ?lat .
  ?s wgs:long ?lng .
  ?category rdfs:label ?category_label .
}
```

Fig. 1. A SPARQL query for extracting tourist attractions

3.1 관광지 및 이벤트 추출

한국관광공사 LOD에서 관광지는 Attraction 클래스에 속하고 이벤트는 Event 클래스에 속한다. Fig. 1과 같은 SPARQL 질의를 사용해서 한국관광공사 LOD로부터 관광지를 추출할 수 있다.

SPARQL 질의의 첫 번째 RDF 트리플 패턴을 보면 rdf:type이 kto:Attraction인 인스턴스를 검색하는 것을 알 수 있다. 여기서 kto는 한국관광공사 LOD에서 기술하고 있는 개체(리소스)에 대한 네임스페이스이고 ktop는 속성에 대한 네임스페이스이다. 이벤트의 경우 같은 SPARQL 질의에서 kto:Attraction을 kto:Event로 변경하여 검색하면 된다. rdfs:label 속성으로 이름을 얻을 수 있고 ktop:category 속성으로 종류를 알 수 있다. 한국관광공사 LOD에는 약 1,900여개의 이벤트와 13,000개의 관광지가 있다.

3.2 위치 추출

관광지 및 이벤트의 위치 정보는 WGS84 어휘(Vocabulary)를 사용한다. WGS84는 위치정보를 표현하는 데에 사용하는 속성들로 구성돼 있다. 한국관광공사 LOD에서는 wgs84_pos/lat와 wgs64_pos/long 속성은 데이터 타입 프로퍼티(DatatypeProperty)로 정의됐으며 위도와 경도를 표현한다. 예를 들어, 한라산 영실의 경우 한국관광공사 LOD에서의 아이디는 126469이므로 위도와 경도는 wgs84_pos/lat와 wgs64_pos/long 속성을 활용하여 다음과 같이 2개의 RDF 트리플로 표현이

된다.

<http://data.visitkorea.or.kr/resource/126469>
 <http://www.w3.org/2003/01/geo/wgs84_pos#lat>
 "33.3402306397"^^<http://www.w3.org/2001/XMLSchema#double> .

<http://data.visitkorea.or.kr/resource/126469>
 <http://www.w3.org/2003/01/geo/wgs84_pos#long>
 "126.4803110551"^^<http://www.w3.org/2001/XMLSchema#double> .

따라서, 관광지 및 이벤트가 열리는 장소의 위도와 경도는 wgs84_pos/lat와 wgs64_pos/long 속성을 가지는 RDF 트리플들로부터 추출할 수 있다.

3.3 내용 추출

관광지 및 이벤트에 대한 설명글을 통해서 내용을 파악한다. 한국관광공사 LOD에서 설명글은 dc:description 속성을 사용한다. 예를 들어, 한라산의 설명글은 다음과 같은 RDF 트리플로 표현이 된다.

<http://data.visitkorea.or.kr/resource/126469>
 <http://purl.org/dc/elements/1.1/description> "제주 한라산은 자연휴식년제를 실시 중인 돈내코 코스를 제외한 다섯 개의 탐방코스가 있다....."@ko .

이벤트의 경우도 마찬가지로 방식대로 설명글을 추출을 한다. 관광지와 이벤트의 설명글을 비교하기 위해 설명글로부터 명사를 추출한다. 문장에서 명사만을 추출하기 위해 한나눔 형태소 분석기를 사용했다[4,5]. 한나눔 태그 셋에서 태그 이름이 ncn으로 시작하는 비서술성 명사와 태그 이름이 nq로 시작하는 고유명사에 해당하는 명사를 추출한다.

3.4 관광지 추천 알고리즘

위와 같은 과정을 거치면 관광지 및 이벤트의 위치(위도와 경도)와 내용(설명글)을 추출할 수 있다. 본 연구에서 이벤트가 열리는 장소 주변의 관광지를 추천한다는 것은 거리가 가까우면서 내용이 유사한 관광지들의 목록을 제시하는 것이다. 관광지 목록을 제시하는 기능은 기

존 관광지 추천 알고리즘과 동일하다. 차이점은 관광지 목록에 포함되는 관광지들을 선별하는 기준이 단일 기준(거리 또는 내용)이 아닌 둘을 모두 고려한다는 점이다. 이를 통해 단일 기준으로 관광지를 선별했을 때의 각각의 단점을 보완하여 좀 더 의미가 있는 관광지 선별이 가능하다.

본 연구에서는 관광지들을 연결하는 도로는 고려하지 않고 단순 직선 거리를 사용한다. Haversine formula를 이용해 위도와 경도를 이용해 두 지점간의 직선 거리를 계산할 수 있다[15].

내용이 얼마나 유사한지를 계산하기 위해 관광지와 이벤트의 설명글로부터 추출된 명사가 얼마나 많이 겹치는지를 본다. 설명글로부터 추출된 모든 명사를 고려하지 않고 자주 사용되면서 특별한 의미를 전달하지 않는 명사들은 제외시킨다. 예를 들어, "다양"과 같은 명사는 대부분의 관광지 및 이벤트의 설명글에 등장하는 단어이지만 특별한 의미를 담고 있지는 않다. 이러한 단어들을 제외시키기 위해 국립국어원의 체언 사전, 행정구역명칭, 한국어 워드넷등을 사용한다[3]. 단순히 명사가 겹치는지만 가지고 판단할 경우 의미적으로 유사한 명사들을 놓칠 수가 있다. 예를 들어, "별자리 관찰"과 관련된 이벤트가 있을 때 이벤트의 설명글에 "천문대"가 명확히 언급되지 않을 수 있다. "별자리" 관련 위키피디아 페이지에 "천문대"가 언급될 경우 "천문대"를 이벤트 설명글에 추가할 수 있다. 이렇게 할 경우 "천문대"를 언급한 관광지와의 유사도를 증가시킬 수 있다. 상기 과정을 알고리즘 형태로 표현하면 Fig. 2.와 같다.

Algorithm: ExtractNouns
Input: Tourist Attraction or Event e, Wikipedia Page Links (P, L), Nouns S
<pre>e.loc = extract_location(e) // Section 3.2 e.con = extract_contents(e) // Section 3.3 e.words = extract_words(e.con) // Extract nouns from the description e.con e.words = remove(e.words, S) // Discard meaningless words e.words = expands(e.words, (P,L)) // Add related words</pre>
Output: Tourist Attraction or Event e with location and contents

Fig. 2. The pseudocode for extracting nouns from description

입력으로 주어져야 하는 위키피디아 페이지 링크 (P, L)은 특정 위키피디아 페이지 P에 포함된 여러 개의 페이지 링크 L를 가리킨다. L에 포함된 페이지 링크 역시

다른 페이지이다. 명사 목록 S는 국립국어원 체언 사전, 행정구역명칭, 한국어 워드넷이다.

주어진 이벤트 e에 대해 유사한 관광지들을 찾아내는 알고리즘은 Fig. 3.과 같다.

Algorithm: TouristAttractionRecommendation
Input: Event e, Korea Tourism Organization LOD O, Wikipedia Page Links (P, L), Nouns S, Weight W
<pre>e = ExtractNouns(e, (P, L), S) attraction_set <- Extract tourist attractions from O using a SPARQL query in Fig. 1. sim_dict = NULL for a in attraction_set a = ExtractNouns(a, (P, L), S) sim = sim(e, a, W) sim_dict.append(e, a, sim) sort(sim_dict)</pre>
Output: sim_dict, a list of tourist attractions sorted by relatedness with event according to location and contents

Fig. 3. The pseudocode for tourist attraction recommendation

우선 ExtractNouns 모듈을 이용해 사용자가 선택한 이벤트의 설명글로부터 명사를 추출한다. 한국관광공사 LOD에서 모든 관광지들을 추출하고 각 관광지에 대해 마찬가지로 ExtractNouns 모듈을 이용해 명사를 추출한다. 주어진 이벤트와 관광지에 대해 명사가 공유되는 정도에 따라 유사도를 계산한다[10,11,13]. 유사도를 계산할 가중치 W도 추가로 입력을 받는데, 0부터 1.0까지의 값으로 위치와 내용에 대한 상대적 중요도이다. 0은 위치만을 고려한 것이고 1.0인 내용만 고려한 것이다. 모든 관광지에 대해 위의 과정을 수행한 뒤 유사도 순으로 정렬하여 반환한다.

추천의 예는 Fig. 4.와 같다. 세종대왕역사문화관(King Sejong The Great History Museum)을 기준으로 관광지 추천을 할 경우, 근처의 관광지로는 세종대왕릉(Tomb of King Sejong), 여주박물관(Yeju Museum), 신륵사(Silleuksa Temple) 등이 있다. 내용이 유사한 관광지는 세종대왕 위키피디아 페이지에 장영실(Jang Yeong-sil) 키워드를 통해서 장영실과학관(Jang Yeong-sil Science Museum)을 찾을 수 있다. 즉, 세종대왕역사문화관의 설명글에는 장영실이 언급이 되지 않아도 위키피디아를 통해 장영실과학관을 찾을 수 있다. 위치와 내용을 모두 고려하면 가중치에 따라 한글 박물관(Hangeul Museum)과 장영실과학관이 상위 순위로 추천이 될 수 있다.

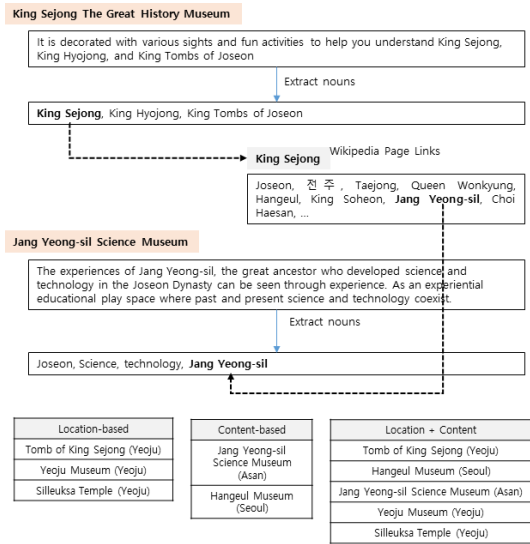


Fig. 4. An example of recommendation

3.5 관광지 추천 예제

Table 1.은 본 연구에서 제시한 관광지 추천 알고리즘에 의해 추천된 예이다. 왼쪽의 이벤트에 대해 오른쪽의 관광지가 추천 순서대로 제시돼 있다. 가중치는 위치와 내용의 상대적 중요도 수치로 TouristAttraction Recommendation 알고리즘에서 W에 해당한다.

Table 1. Examples of tourist attraction recommendation

Event	Recommended tourist attractions		
	weight		
	0	0.5	1.0
Jungeupsa cultural festival	Beopinsa	Gomanaru Ferry	Cheoneunsa
	Two Stone Standing Buddhas	Sudeoksa	Sudeoksa
	Ibamsan	Cheoneupsung	Cheoneupsung
Yulgok cultural festival	Hwaseokjeong	Hwaseokjeong	Obong Confucian Academ
	Tomb of Song Si-yeol	Obong Confucian Academy	Daeseongjeon Shrine of Mujang Local Confucian School
	Hwayangguk	Tomb of Song Si-yeol	Yongam Local Confucian School

정읍사문화제(Jungeupsa cultural festival)에 대해 가중치가 0인 경우 정읍사문화제가 열리는 정읍사문화공원 주변의 법인사(Beopinsa), 석조이불입상(Two Stone

Standing Buddhas), 입암산(Ibamsan) 등이 추천됐다. 입암산의 경우 정읍사문화제의 내용과 직접적인 관련은 없다. 가중치가 1.0인 경우 천은사(Cheoneunsa), 수덕사(Sudeoksa), 천읍성(Cheoneupsung)이 추천됐다. 천은사와 천읍성의 경우 충렬왕과 관련된 사찰인데, 정읍사문화제의 “정읍사” 단어와 “충렬왕” 단어가 위키피디아에 의해 연결이 되었기 때문에 추천되었다. 수덕사의 경우 백제의 사찰인데 정읍사가 백제의 사찰이기 때문이다. 하지만, 천은사의 경우 정읍사문화제가 열리는 장소와 거리상 멀리 떨어져 있어서 정읍사문화제에 참여한 관광객에게 추천하는 것은 적절하지 않다. 이와 같이 위치기반 추천(가중치 0)과 내용기반 추천(가중치 1.0)은 부적절한 추천이 포함돼 있다. 본 연구에서 제안하는 방법인 위치와 내용을 모두 고려하는 경우(가중치 0.5)에서는 대신에 고마나루가 추천되었다. 고마나루는 백제 역사의 중심이자 국제적 교통의 관문으로 정읍사문화제가 열리는 장소에서 다른 관광지에 비해서 상대적으로 가까운 곳이다. 천은사를 추천하는 것보다는 고마나루를 추천하는 것이 정읍사문화제에 참여하고 있는 관광객에게 더 도움이 될 것이다.

울곡문화제(Yulgok cultural festival)는 파주이이유적지, 자운서원 등의 장소에서 개최됐다. 인근에 있는 화석정(Hwaseokjeong), 송시열묘(Tomb of Song Si-yeol), 화양구곡(Hwayanggugok) 등의 관광지가 추천됐다. 가중치 1.0인 경우 울곡 이이와 송시열이 연관이 많기 때문에 송시열의 영정이 모셔져 있는 오봉서원(Obong Confucian Academy)이 추천됐다. 하지만 오봉서원은 화석정에 비해 먼 곳에 위치해 있기 때문에 가중치 0.5의 경우와 같이 오봉서원보다는 화석정을 우선해서 추천하는 것이 울곡문화제에 참여한 관광객에게 더 도움이 될 것이다.

4. 결론

본 논문에서는 위치와 내용을 모두 고려한 관광지 추천 알고리즘을 제안했다. 한국관광공사 LOD, 용어 사전, 위키피디아 등에 기반하여 이벤트와 관광지 간의 내용 유사도를 계산하였고 위도/경도 기반으로 거리를 합산하여 관광지를 추천하는 방법을 고안했다.

학문적인 관점에서는 컴퓨터공학 분야의 텍스트마이닝 기술을 관광학 분야에 적용했다는 점에서 과학기술과 인문사회 분야의 융합 연구로서 의미를 찾을 수 있다. 본

연구에서는 위키피디아를 기반 데이터로 사용했는데 블로그 등의 사용자의 견해가 실시간으로 반영되는 매체를 활용하게 되면 좀 더 의미있는 추천이 가능할 수 있다. 이를 위해서 실시간으로 블로그의 글을 크롤링하는 기술을 융합하는 연구가 가능하다.

또한, 기존 대부분의 온톨로지 기반 관광지 추천 연구의 경우 사용자 프로파일을 온톨로지로 구성하여 사용자의 취향을 분석한 뒤 관광지를 추천하는 데에 활용한 반면 본 연구에서는 사용자 프로파일이나 관광지의 설명글에 대한 텍스트마이닝을 했다는 점에서 기존 연구와 차별점을 찾을 수 있다.

실무적인 관점에서, 본 연구는 카카오맵 등의 지도 서비스에 추가 기능으로 활용될 수 있다는 점에서 의의를 찾을 수 있다. 관광객이 관광 도중에 자주 사용하는 지도 서비스에 본 연구의 아이디어가 쉽게 적용될 수 있다는 점에서 실무적인 의미가 크다고 할 수 있다. 특히, 기반 데이터로 온톨로지를 활용했는데 온톨로지는 표준화된 데이터 표현방식이기 때문에 웹상에 공개된 관광 관련 최신 데이터를 쉽게 연동할 수 있는 장점도 있다.

향후 연구에서는 위도/경도 기반의 단순 직선 거리가 아닌 연결된 도로를 고려한 거리를 고려하도록 개선할 예정이다. 또한, 실제 사용자를 대상으로 한 실험을 통해 위치와 내용에 대한 가중치를 정하는 방법에 대한 연구를 통해 보다 사용자 맞춤형 관광지 추천 시스템이 되도록 개선할 예정이다.

References

- [1] J. Borràs, A. Moreno, A. Valls, "Intelligent tourism recommender systems: A survey" *Expert Systems with Applications*, Vol.41, Issue.16, pp.7370-7389, 2014
DOI: <https://doi.org/10.1016/j.eswa.2014.06.007>
- [2] M. S. Kim, H. M. Um, "A Comparison of First Time and Repeat Visitors' Tourism Destination - Focusing on Seoul City", *Korea Academy Industrial Cooperation Society*, Vol.17, No.11, pp.648-654, 2016
DOI: <https://doi.org/10.5762/KAIS.2016.17.11.648>
- [3] J. H. Ahn, E. H. Kim, H. G. Kim, "Big Data based Tourist Attractions Recommendation - Focus on Korean Tourism Organization Linked Open Data", *Management & information systems review*, Vol.36, No.4, pp.129-148, 2017
DOI: <https://doi.org/10.29214/damis.2017.36.4.008>
- [4] S. W. Park, D. H. Choi, E. K. Kim, K. S. Choi "A Plug-In Component-based Korean Morphological Analyzer", *The 22th Annual Conference on Human &*

- Cognitive Language Technology*, pp.197-201, 2010
- [5] A. S. Yoon, S. H. Hwang, E. R. Lee, H. C. Kwon, "Construction of Korean Wordnet KorLex 1.5", *Korea Information Science Society*, Vol.36, No.1, pp.92-108, 2009
- [6] S. W. Byun, "A study on the effect of the motivation of choosing the tourist spots on the image of the tourist spots and the intention of revisiting - Focusing on Gamcheon Culture Village", *Management & information systems review*, Vol.34, No.3, pp.197-213, 2015
DOI: <https://doi.org/10.29214/damis.2015.34.3.011>
- [7] S. R. Kang, J. H. Bang, M. S. Kim, "An Exploratory Study on Foreign Inbound Tourist in the Tourism Industry: expenditure, satisfaction, and intention to revisit", *Journal of the Korea Academia-Industrial cooperation Society*, Vol.18, No.9, pp.315-321, 2017
DOI: <https://doi.org/10.5762/KAIS.2017.18.9.315>
- [8] T. Y. Kim, Y. M. Hwang, S. A. Cho, "Study on Travel Path of Tourists -Focus on Gyeongnam Tourists", *Journal of Tourism Studies*, Vol.22, No.2, pp.247-268, 2010
- [9] Y. M. Choi, R. J. Hong, "The Effect of Online Tourist Information on Selection of Cultural Heritage Destination: In Four Royal Palaces", *Journal of Culture Industry*, Vol.17, No.4, pp.9-17, 2017
- [10] B. B. Cambazoglu, E. Kayaaslan, S. Jonassen, C. Aykanat, "A term-based inverted index partitioning model for efficient distributed query processing", *ACM Trans. Web*, Vol.7, No.3, pp.1-23, 2013
DOI: <https://doi.org/10.1145/2516633.2516637>
- [11] K. Chen, Z. Zhang, J. Long, H. Zhang, "Turning from TF-IDF to TF-IGM for term weighting in text classification", *Expert Systems with Applications*, Vol.66, No.C, pp.245-2600, 2016
DOI: <https://doi.org/10.1016/j.eswa.2016.09.009>
- [12] C. Bizer, T. Heath, and T. B. Lee, "Linked data-the story so far", *International Journal on Semantic Web and Information Systems*, Vol.5, No.3, pp.1-22, 2009
DOI: <http://dx.doi.org/10.4018/jswis.2009081901>
- [13] J. H. Paik, "A novel TF-IDF weighting scheme for effective ranking", *Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval*, pp.343-352, 2013
DOI: <https://doi.org/10.1145/2484028.2484070>
- [14] L. Yu, "Linked open data: A Developer's Guide to the Semantic Web", *Springer Berlin Heidelberg*, pp.409-466, 2011
DOI: <https://doi.org/10.1007/978-3-642-15970-1>
- [15] V. Brummelen, G. Robert, "Heavenly Mathematics: The Forgotten Art of Spherical Trigonometry." Princeton University Press, 2013
DOI: <https://doi.org/10.1080/10848770.2015.1028011>
- [16] Y. J. Park, K. A. Song, J. W. Whang, B. M. Chang, "A System for Personalized Tour Recommendation Based on Ontology", *Journal of the Korea Contents Association*, Vol.15, No.9, pp.1-10, 2015
DOI: <https://doi.org/10.5392/JKCA.2015.15.09.001>
- [17] T. Majeed, E. A. E. Stämpfli, A. Liebrich, R. Meier, "Personalized Hybrid Recommendations for Daily Activities in a Tourist Destination", *The 9th International Symposium on Ambient Intelligence*, Toledo, Spain, pp.155-165, June 2018.
- [18] Z. Bahramian, R. A. Abbaspour, C. Claramunt, "Toward Geospatial Collaborative Tourism Recommender Systems." *GIS Applications in the Tourism and Hospitality Industry*, pp.212-248, 2018
DOI: <https://doi.org/10.4018/978-1-5225-5088-4.ch010>

안진현(Jinhyun Ahn)

[정회원]



- 2005년 8월 : 고려대학교 컴퓨터교육학과 (이학사)
- 2007년 8월 : 고려대학교 컴퓨터교육학과 (이학석사)
- 2017년 2월 : 서울대학교 의료경영과정보학 (공학박사)

- 2017년 3월 ~ 2018년 2월 : 서울대학교 의생명지식공학 연구실 선임연구원
- 2018년 3월 ~ 현재 : 제주대학교 경영정보학과 조교수

〈관심분야〉

분산/병렬 컴퓨팅, 지식공학, 비식별화, 온톨로지

임동혁(Dong-Hyuk Im)

[정회원]



- 2003년 2월 : 고려대학교 컴퓨터교육학과 (이학사)
- 2005년 2월 : 서울대학교 컴퓨터공학과 (공학석사)
- 2011년 2월 : 서울대학교 컴퓨터공학과 (공학박사)

- 2011년 3월 ~ 2012년 8월 : 서울대학교 의생명지식공학 연구실 선임연구원
- 2013년 3월 ~ 현재 : 호서대학교 컴퓨터공학과 부교수

〈관심분야〉

빅데이터 컴퓨팅, 비식별화, 그래프 데이터베이스