

사용자 프로파일과 상황정보를 이용한 여행지 및 방문경로 추천시스템

김영민* · 변영철** · 이상준**

Tour-place and Path Recommendation System using User Profile and Context Information

Young-Min Kim*, Yung-Cheol Byun** and Sang-Joon Lee**

ABSTRACT

In this study we design and implement tour information recommendation and path scheduling system which recommends the places to visit and guides a path for a traveler who wants an information in a strange place. System gets user profile information to recommend some preferable places for the user and also it collects context informations, for example weather, season. And then the system classifies the user and calculates which places are preferable for the user. After the user selects places to visit from the list of recommendation, finally the system shows a reasonable path to visit sequentially. Through this experimentation, we found that this recommending system can work as an human expert to guide tour informations.

Key Words : Recommendation system, profile, context-aware, path scheduling

1. 서론

여행자에게 있어서 낯선 여행지를 여행할 때 가장 필요한 것은 잘 정리되어 있고 최신 정보로 꾸며진 안내책자이거나 현지 여행지 정보에 밝은 한 사람의 가이드일 것이다. 안내 책자는 사용자가 자신의 선호도에 따라서 여행지를 선택할 수 있도록 모든 정보를 제공할 것이며, 가이드는 여행자의 선호도와 현지의 상황정보 등을 고려하

여 최적의 여행지를 소개하고 안내해 줄 것이다. 안내 책자에 비해 가이드가 좀더 동적이고 현실적이며 그날의 날씨나 도로상황, 현지 여행지 상황정보등을 고려하여 최적의 관광지 추천을 할지도 모른다. 본 논문은 이와같은 한사람의 현지 정보에 밝은 가이드를 모방한 여행지 추천 및 방문 경로 안내 시스템을 제안한다.

기존의 차 등[1]의 논문은 단순히 모바일 단말기 상에서 관심 여행지를 탐색할 수 있는 수준에서 그친다. 고 등[2]은 온톨로지를 이용하여 여행지 검색의 정확도를 높였지만 여전히 사용자가 직접 검색해야 한다. Cheverst[3], Abowd[4], Yang[5]의 논문은 사용자의 검색이 아닌 사용자 위치를 시스템이 파악하여 해당 위치에 해당하는 여행지 정보를 제공하는 상황인지(context-aware) 개념을 도입하였다. 하지만 단순히 위치만을 이

* 제주대학교 대학원 정보공학과
Dept. of Information Engineering, Cheju National University
** 제주대학교 통신컴퓨터공학부, 첨단기술 연구소
Faculty of Telecommunication and Computer Eng.,
Research Institute of Advanced Technology, Cheju
National University

용한 단순 서비스에 그쳤으며 Pashtan[6]의 경우는 여행자가 차량 이동중일때 그 속도와 방향, 그리고 시간정보까지도 컨텍스트(context)정보로 활용한 여행지 안내를 수행한다. 본 논문은 기계학습을 이용한 지능형 시스템의 대표적인 방법인 뉴럴 네트워크(neural network)를 이용해 사용자의 프로파일 정보와 웹 서비스를 이용한 날씨정보 등의 상황정보에 맞는 여행지 그룹을 추천하는 시스템을 제안한다[7,8,9,10].

또한 추천된 관광지 그룹들 중에서 사용자가 선호지역을 선택하면 이들의 방문순서를 계산해 준다. 임[11]과 Liu[12]는 도로 정보 및 공공 교통 시스템을 고려한 방문순서를 제안하였으나 이들 시스템은 도로 기준이 아닌 수많은 관광지를 방문지를 기준으로 한 연산인 경우는 연산량이 너무 많아 문제 해결이 쉽지 않다. 이러한 문제들을 해결하기 위해 최근 등장한 유전자 알고리즘(Genetic Algorithm)은 이러한 방문 순서등의 문제 해결에 적합하다[13]. 따라서 본 논문은 이러한 GA를 이용해 최적의 방문지 순서를 생성하고 추천함으로써 여행자가 주어진 시간안에 최적의 순서로 여행지를 방문할 수 있는 편의를 제공하게 된다.

본 논문의 구성은 2장에서 사용자 프로파일 구성 및 신경망을 이용한 추천시스템과 유전자 알고리즘을 이용한 최적방문경로 탐색 방법에 대해 소개한다. 3장에서 이를 구현 및 테스트하고 4장에서 결과를 논한다.

II. 제안하는 시스템

여행을 원하는 사용자가 기본적인 프로파일 정보를 입력하면 시스템은 해당 사용자에게 적합한 관광지를 검색하여 추천해주고 사용자는 자신이 원하는 선호 지역을 선택하여 스케줄링을 요구하게 된다. 시스템은 사용자가 선택한 지역에 사용자 프로파일 정보를 합하여 최적의 여행 스케줄링을 수행한다.

시스템의 추천 지역 외에 사용자가 추가로 원하는 관광지역이 있다면 이것 또한 해당 사용자 프로파일과 관련된 관심지역이 될 수 있음으로

feedback을 수행하여 재 학습하도록 한다. 사용자의 여행 일정이 당일 여행일수도 있고 1박 2일 내지 1주일일 수도 있으며 이에 따른 최적 계획을 제안할 수 있어야 한다. 또한 사용자 프로파일인 경우 가족동반 여행, 단체여행, 소수여행, 계절별 테마 여행 등등 다양한 경우에 따라 구별된 추천을 할 수 있어야 하며, 날씨와 그 해 또는 지역별 특별 행사등 동적으로 변할 수 있는 정보들을 감안하여 복잡한 경우를 충족하는 추천 및 스케줄링이 가능해야 한다

2.1 시스템 구조

전체 시스템 흐름도는 다음 Fig. 1과 같다. 방문횟수 및 단체여행 여부, 가족동반여부, 자가 운전여부, 선호도 등의 사용자 프로파일은 신경망 입력으로 들어가며 사용자 입력 외에 날씨와 계절 등의 동적정보가 신경망 입력이 된다. 신경망 시스템은 사용자 프로파일과 동적정보를 기반으로 최적의 여행지를 추천하며 이때 여행지 관련 정보를 DB에서 검색하여 자세한 설명도 함께 보여준다. 사용자는 시스템이 추천한 여행지 중에서도 특히 선호하는 지역과 비선호 지역을 체크하여 스케줄링을 하도록 재차 시스템에 요구하며 시스템은 유전자 알고리즘이 제시하는 최적의 방문경로를 생성하여 사용자에게 제시하여 준다.

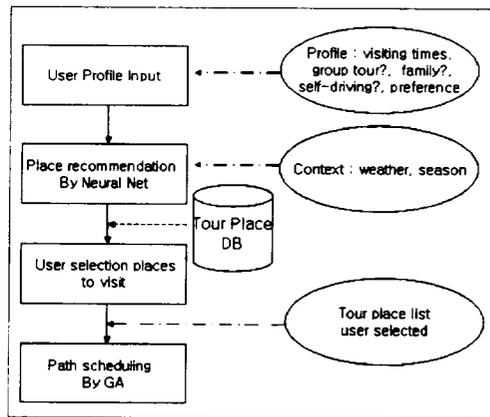


Fig. 1 System workflow

전체 시스템 구조도는 다음 Fig. 2와 같다. 크게 사용자 인터페이스 부분과 전처리

(Preprocessor), 여행지 선택기(Tour-place Selector), 관광지 추천 및 피드백 에이전트(Tour-place recommend and feedback agent), 상황정보 수집기(Context information collector) 그리고 마지막으로 방문경로 스케줄러(Path scheduler)로 구성된다.

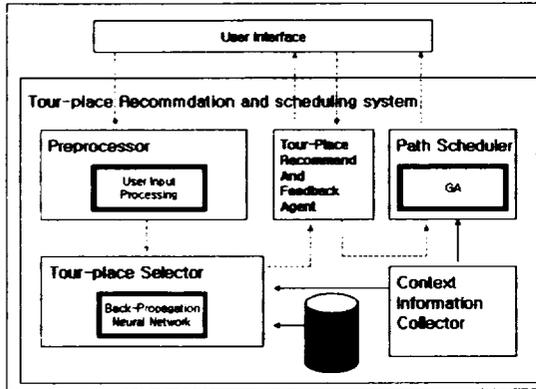


Fig. 2 System architecture

전처리는 사용자의 입력 프로파일값을 신경망 입력으로 사용하기 위해 특정범위의 수치값으로 스케일링하고 입력패턴을 생성해 주는 역할을 담당한다. 여행지 선택기는 역전파 뉴럴 네트워크(back propagation neural network)에 의해 학습된 상태에서 사용자 입력 프로파일 패턴이 들어오면 동적정보 패턴과 함께 학습된 결과대로 사용자 프로파일에 적합한 여행지 그룹을 추천해주는 부분을 담당한다. 여행지 추천 및 피드백 에이전트는 여행지 선택기가 추천한 여행지 그룹에 해당하는 여행지들의 상세한 정보들을 사용자에게 제공하며, 선호지역과 비 선호지역을 입력받아 재 정렬한 후 여행 스케줄러에게 넘겨주는 역할을 수행한다. 동적정보 수집기는 웹 서비스(web service)에 의해 다양한 정보 제공 서버로부터 정보를 수집하며 날씨정보와 교통정보, 계절 등의 정보를 수집하여 신경망 선택기에 제공한다. 경로 선택기는 여행경로 생성문제에 가장 적합한 해결기법인 유전자 알고리즘(Genetic Algorithm)에 의해 사용자 선호 지역에 대한 방문 순서를 생성하여 준다.

2.2 신경망 입력을 위한 사용자 프로파일 구성

신경망 입력 패턴을 얻기 위해서 사용자 인터페이스로부터 다음과 같은 입력을 얻게 된다.

1. 첫 방문이십니까?	예/아니오
2. 가족동반이십니까?	예/아니오
3. 자가운전을 하십니까?	예/아니오
4. 단체여행을 하고계십니까?	예/아니오
5. 체험관광을 좋아하십니까?	예/아니오
6. 전시 관람 및 공연을 좋아하나요?	예/아니오
7. 현재 날씨는?	맑음/흐림
8. 여름인가요?	예/아니오

Fig. 3 User input form values

0과 1의 비트패턴 생성을 위해 입력은 '예'와 '아니오'를 선택받도록 하였으며 현재 날씨는 '흐림'을 '0'으로 '맑음'을 '1'의 형태로 처리하게 된다. 날씨와 계절은 동적정보 처리기에 의해 수집되어 만들어지는 정보이다. 예를 들어 첫 방문이며 가족동반여행이고, 자가운전은 하지 않고 단체여행도 아니며 체험관광을 좋아하고 공연을 좋아할 때, 날씨는 좋고 여름이 아니면 '11001110'과 같은 8개의 비트 패턴이 생성된다.

2.3 사용자 Profile에 따른 추천 및 선호 지역 학습

신경망을 이용한 학습시스템은 BPN(back-propagation network)를 이용해 학습시키며 입력으로 8비트, 출력으로 18비트를 사용한다. 사용자가 작성한 프로파일 정보와 동적정보값을 비트패턴화한 8비트의 입력 패턴에 따라, 각각 어떠한 관광지 그룹 패턴을 출력해야 하는지 각각의 입력 패턴마다 추천 지역을 학습시킨다. 학습을 위해 현재 대부분의 제주 여행관련 사이트에서 관광지 분류를 위해 사용하는 18개의 카테고리를 사용하여 출력 패턴을 구성하였다. 즉 시스템은 사용자 프로파일에 따라 18가지의 관광지 타입에 대한 선호 경향을 학습하게 된다. 다음 Table 1은 관광지 추천을 위한 관광지 그룹의 18비트 이진 스트링 표현이다.

Table 1 BPN output bit pattern

no	bit	group type
1	0	valley
2	0	cave
3	0	island
4	0	underwater tour
5	0	indoor performance
6	1	indoor exhibition
7	1	indoor experiment
8	0	outdoor performance
9	0	outdoor sightseeing
10	0	outdoor exhibition
11	0	outdoor experiment
12	0	outdoor resorting
13	0	climbing
14	0	boat pleasure
15	0	historic place
16	0	water falls
17	0	beach
18	0	coastal landscape

Table 1이 시스템 출력으로서 표현하는 비트 패턴은 '000001100000000000'이며 그 의미는 실내 전시, 실내체험에 관심이 있다는 의미가 된다.

2.4 사용자 선호 지역 순회를 위한 방문 스케줄링

사용자가 선택한 선호 지역 정보들은 유전자 알고리즘(genetic algorithm)의 입력으로 들어가며, 유전자 알고리즘은 이를 비트스트링 형태의 유전자 정보를 생성하고 이에 대한 최적의 유전자 배열을 만들어 낸다. 최적의 유전자 배열은 곧 최적의 여행 비용을 의미하며 시스템은 이를 최고의 스케줄링 결과라 판단하고 사용자에게 제공한다.

GA 연산을 위해 사용한 정보는 다음과 같다. 유전정보의 표현은 2진 스트링을 사용했으며(이는 여행스케줄링의 일반적인 유전정보 표현 형태), 교배확률(crossover), 돌연변이 확률(mutation)은 각각 0.25와 0.05로 했으며 개체수(population size)는 스케줄링 지역의 2배수, 진화 세대수(evolution time)은 스케줄링 대상 지역수의 10배를 한계로 하였다.

Table 2 Genetic algorithm parameters

Crossover : 0.25%
Mutation : 0.05
Population Size : number of target regions * 2
Evolution times : number of target regions * 10

GA에서의 각 유전자에 대한 평가는 <http://mincando.pe.kr/ontology/JejuRegions.owl>에 기술된 제주지역 온톨로지의 각 지역간 인접정보를 기준으로 평가를 수행하였다. 다음 그림은 지역간 인접정보를 표현한 행정구역 온톨로지의 일부이다.

```

<kr:Dong
rdf:about="http://mincando.pe.kr/ontology/KoreaRegions.owl#R3901057">
  <kr:code>3901057</kr:code>
  <kr:name xmi:lang="ko">용담1동</kr:name>
  <kr:ename>Yongdam1(il)-dong</kr:ename>
  <locatedIn>
    <kr:Si
rdf:about="http://mincando.pe.kr/ontology/KoreaRegions.owl#R39010">
  </locatedIn>
  <borderRegion>
    <kr:Dong
rdf:about="http://mincando.pe.kr/ontology/KoreaRegions.owl#R3901058">
  </borderRegion>
  <borderRegion>
    <kr:Dong
rdf:about="http://mincando.pe.kr/ontology/KoreaRegions.owl#R3901055">
  </borderRegion>
</kr:Dong>
    
```

Fig. 4 part of jeju region ontology

이 온톨로지는 기본적으로 각 지역이 위치한 상위 구역과 각 지역이 인접한 지역정보들을 표현해주고 있다. 이러한 인접관계를 추적하여 각 지역간 거리비용을 계산할 수 있다.

현재로서는 단순히 인접깊이 D(distance)로

$$D = \text{getDepthLevel}(\text{fromRegion}, \text{toRegion}) \quad (1)$$

그 평가 점수를 계산하였다. 즉, 관광지들이 속한 지역정보들을 가지고 지역정보의 인접 정도가 멀수록 비용이 크다고 보고, 전체 인접 정도가 가장 짧은 여행지 순서를 최적의 방문순서로 결정하도록 하였다.

III. 구현 및 테스트

3.1 시스템 환경 및 데이터 구성

전체 시스템 구현 환경은 다음과 같다.

Language : Java J2SDK1.4.2_02

Server : Tomcat 5.0.8

테스트 데이터의 구성은 제주도 관련 249개 관광지들을 대상으로 18개의 관광지 그룹을 구성하여 추천하도록 하였다.

3.2 사용자 프로파일 및 상황에 따른 학습

Fig. 5는 BPN에서 사용한 학습 데이터의 일부를 보여주고 있다. 학습데이터는 8비트의 입력패턴과 18비트의 출력패턴으로 이루어진 학습패턴 256개에 대하여 학습을 수행한다.

256 <== number of data	00100101 10000000000010011
8 <== input bits	00100110 00000010000000000
18 <== output bits	00100111 00000010000000000
00000000 0000011000000000000	00101000 00000110000000000
00000001 00000001001100001	00101001 00001010000000000
00000010 00000001001100001	00101010 00001100000000000
00000011 00000001001100001	00101011 00001010000000000
00000100 00000001000100100	00101100 00000001000100100
00000101 10000000000010011	00101101 00000000000000010
00000110 0000010001000000000	00101110 00000001000000000
111110101 00000001000000101	00101111 00000001000000000
	11111011 0101111000000000010

Fig. 5 Training data for BPN

3.3 사용자 입력에 따른 관광지 추천

Fig. 6은 여행지 추천을 위해 사용자 프로파일 정보를 입력받는 인터페이스 화면이다. 전체 8개의 프로파일중 6개는 사용자가 입력하며 마지막 두 개는 웹 서비스(web service)를 이용해 처리한다.

1. 첫 방문이십니까?	<input checked="" type="radio"/> 예 <input type="radio"/> 아니오
2. 가족동반이십니까?	<input type="radio"/> 예 <input checked="" type="radio"/> 아니오
3. 자가운전을 하실겁니까?	<input type="radio"/> 예 <input checked="" type="radio"/> 아니오
4. 단체여행을 하고계십니까?	<input type="radio"/> 예 <input checked="" type="radio"/> 아니오
5. 체험관광을 좋아하십니까?	<input type="radio"/> 예 <input checked="" type="radio"/> 아니오
6. 현재 날씨는?(Web Service로 대처)	<input checked="" type="radio"/> 맑음 <input type="radio"/> 흐림
7. 전시 관람 및 공연을 좋아하나요?	<input type="radio"/> 예 <input checked="" type="radio"/> 아니오
8. 여흥인가요(계절 정보에 대한 Web Service로 대처)	<input type="radio"/> 예 <input checked="" type="radio"/> 아니오

시작

Fig. 6 User interface for profile input

이렇게 입력된 사용자 프로파일 정보와 웹 서비스를 이용해 동적으로 수집된 상황정보(context information)은 8비트의 입력 패턴을 구성하게 되

며, 이에 대해 학습된 패턴을 출력하게 되면 다음 Fig. 7과 같은 추천된 여행지 그룹별로 선호 지역을 선택할 수 있도록 추천을 수행한다.

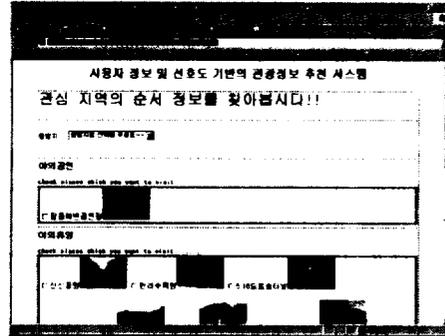


Fig. 7 Four-place recommendation

Fig. 7에서 보면 야외공연 그룹과 야외 휴양 그룹이 추천되어진 것을 볼 수 있다. 이런 경우는 동적정보가 제공하는 날씨 정보가 맑은 경우이다. 제일 위 부분에는 선호지역을 선택했을 경우에 방문을 위한 출발지를 선택해주는 입력이 필요하다.

사용자의 선호지역이 비자림, 제주성지, 탐동해변공원장, 함릉유적지, 단산이 선택 되었을 경우에 제주의 동쪽 지역부터 출발하여 제주시 지역, 다음으로 제주 서부지역 순으로 방문순서를 추천하고 있다. Fig. 8은 보면 '비자림' => '제주성지' => '탐동해변공원장' => '함릉유적지' => '단산' 임을 알 수 있다.

요청지역 :

- 출발지 : 봉개동
- 방문지 : 탐동해변공원장, 비자림, 단산, 제주성지, 함릉유적지

추천 방문 순서입니다!!

- Using Region Ontology of Jeju : <http://mincando.pe.kr/ontology/jeju>
- Using Region Ontology of Korea : <http://mincando.pe.kr/ontology/kr>
- Using Genetic Algorithm for optimizing the path list to visit
- Crossover : 0.25%, Mutation : 0.05, Population Size : number of tour

10

The Optimized(Proper???) Path List

비자림 => 제주성지 => 탐동해변공원장 => 함릉유적지 => 단산

Fig. 8 Result of path scheduling

IV. 결 론

낮선 여행지를 방문하는 여행자에게 있어서 길을 안내해주는 가이드가 필수적이다. 인간 가이드와 같이 사용자의 정보에 맞는 방문지를 추천해주는 시스템은 그만큼 여행자들에게는 큰 도움이 될 것이다.

본 논문은 이와 같은 추천시스템을 위해 사용자 프로파일을 분석하고 이에 맞는 여행지를 추천해 주기 위해 BPN을 이용하였으며 효과적으로 관광지 그룹을 선정해 줌을 보였다. 또한 사용자가 선택한 선호 지역 정보들은 유전자 알고리즘(GA)의 입력으로 들어가며, 유전자 알고리즘에 의해 최적의 유전자 배열인 여행지 방문 경로를 만들어 줌을 볼 수 있었다.

참고문헌

- 1) 차종우, 김현수, 안철준, 조미경, 2003, PDA기반의 부산경남 문화관광 안내 시스템의 설계 및 구현, 정보처리학회 추계학술발표, 제 10권 제 2호, pp 837-840
- 2) 고은정, 김여정, 김운, 강지훈, 2004, 지역 온톨로지를 이용한 지능형 여행정보 제공 시스템, 정보과학회 2004년 춘계학술대회, pp 610-612
- 3) K. Cheverst et al., 2000, Experiences of Developing and Deploying a Context-Aware Tourist Guide: The GUIDE Project, Proc. MOBICOM 2000: 6th Ann. Int'l Conf. Mobile Computing and Networking, ACM Press, NewYork, 2000, pp. 20-31.
- 4) Gregory D. Abowd, Christopher G. Atkeson, Jason Hong, Sue Long, Rob Kooper and Mike Pinkerton, 1997, Cyberguide: a mobile context-aware tour guide, Special issue: mobile computing and networking: selected papers from MobiCom '96, Vol 3, pp 421-433
- 5) Jie Yang, Weiyi Yang, Denecke M. and Waibel A., 1999, Smart sight:a tourist assistant system, 3rd International Symposium on Wearable Computers, pp 146-153
- 6) A. Pashtan, A. Heusser, P. Scheuermann and R.

- Blaettler, 2003, A Context-Aware Tourist Information System, Proceedings 4th International Workshop of Mobile Computing.
- 7) Ingo Schwab and Wolfgang Pohl, 1999, Learning User Profiles from Positive Examples, In Proceedings of the International Conference on Machine Learning & Applications, pp. 15-20
- 8) TingFan Wu, Chih-Jen Lin, Ruby C. Weng, 2004, Probability Estimates for Multi-class Classification by Pairwise Coupling, Journal of Machine Learning Research, 5, pp 975-1005
- 9) G. L. Foresti, C. Micheloni, 2002, Generalized neural trees for pattern classification, IEEE Trans on Neural Networks, Vol 13, No 6, pp 1540-1547
- 10) Neural Net in Java
<http://diwww.epfl.ch/mantra/tutorial/english/>
- 11) 임재걸, 1998, 휴리스틱 알고리즘을 사용하여 국토 이용 여행 최적 경로를 제공하는 웹 사이트 구현, 98년 정보처리학회 추계학술대회, pp 381-384
- 12) Chao-Lin, Tun-Wen Pai, Chun-Tien Chang and Chang-Ming Hsieh, 2001, Path-Planning Algorithms for public Transportation Systems, The 4th Internation IEEE Conference on Intelligent Transportation Systems
- 13) 공성근, 김인택, 박대회, 박주영, 신요안, 1996, 유전자 알고리즘, 그린출판사
- 14) <http://mincando.pe.kr/ontology/JejuRegions.owl>
- 15) Jena - a semantic web framework for java
<http://jena.sourceforge.net/>