

의미망과 베이지안 네트워크 기반 모바일 라이프 브라우저 개발

황금성, 이종욱, 이영설, 조성배
연세대학교 컴퓨터과학과

yellowg@sclab.yonsei.ac.kr, mcchun0@hotmail.com, tiras@sclab.yonsei.ac.kr, sbcho@cs.yonsei.ac.kr

Development of Mobile Life Browser based on Semantic and Bayesian Network

Keum-Sung Hwang, Jong-Wook Lee, Young-Seol Lee, Sung-Bae Cho
Dept. of Computer Science, Yonsei University

1. 서론

개인화 장비의 기술이 발전하면서 다양한 로그 정보 수집이 가능해졌고, 이를 통한 사용자 정보 수집이 가능해지고 있다. 최근에는 이러한 로그, 즉 모바일 환경에서 수집한 개인의 라이프 로그를 분석하여 일상생활 기록으로 남기거나 분석하여 사용자에게 서비스를 제공하려는 연구가 활발히 진행되고 있다. 특히, 라이프 브라우저는 마이크로소프트 연구소의 E. Horvitz가 처음으로 사용한 용어로서 당시에는 개인과 관련한 정보를 효과적으로 검색할 수 있도록 지원하는 도구를 지칭하였다[1].

본 논문에서는 모바일 디바이스로부터 수집된 정보를 데이터베이스화하고 이를 분석하여 사용자의 일상을 요약한 후 효율적으로 시각화하기 위한 라이프 브라우저를 연구하고 개발한다. 라이프 브라우저는 개인의 일상을 보다 효과적으로 관리할 수 있도록 해 주며, 일상 로그를 활용한 다양한 부가 서비스들을 손쉽게 개발할 수 있도록 해줄 수 있을 것이다.

1.1 이전 연구

앞선 연구[2]에서 저자들은 이미 모바일 라이프 로그를 분석해서 모바일 라이프 브라우저 형태로 표현하고, 사용자 생활 요약 및 검색 서비스를 지원하는 연구를 수행하였다. 본 논문은 앞선 연구[2]를 바탕으로 확장 진행된 연구이며, 오픈 소스 API의 매쉬업과 검색을 위한 의미 기반 검색 알고리즘이 추가되었다. 본 논문에서는 MyLifeBrowser 라는 이름으로 개발되었으며, 모바일 디바이스 사용자에게 의해 수집된 데이터를 관찰하고 검색할 수 있다.

2. 모바일 라이프 브라우저

그림 3은 본 논문에서 개발된 모바일 라이프 브라우저의 전체적인 시스템을 나타낸다. 모바일 디바이스에서 로그 데이터 수집이 이뤄지면 이를 DB 서버로 전송한 뒤, DB를 중심으로 전체적인 시스템이 동작하게 된다. 웹 환경에서 동작하도록 개발하기 위해 이를 지원하는 서버 기능 및 오픈소스 API가 사용되었다.

2.1 라이프 로그 분석

모바일 디바이스로부터 수집된 정보는 전처리를 통해서 분석하기 좋은 상태로 가공되고, 다시 확률 모델에

의해 특이성으로 추출된다. 본 논문에서는 기존 연구 [2][3]에서 수행된 데이터 이산화(discretization), 통계적 분석, 임팩트 분석, GPS 위치 분석, GPS 오류 처리, 결손 데이터 추산, 확률 추론 과정을 그대로 수행하였다.

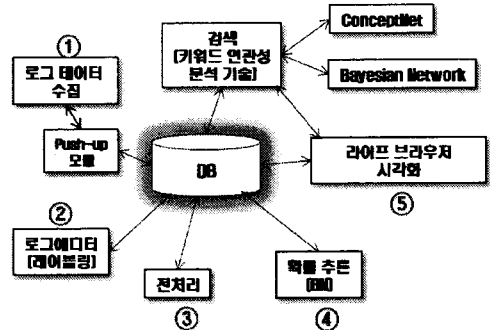


그림 3. 라이프 브라우저 구현을 위한 시스템 구조도.
기본적인 정보 흐름은 ①~⑤와 같다.

2.2 의미 기반 검색

MIT Media 연구실에서 개발한 ConceptNet은 상식적 자연어 형식으로 가공한 Toolkit으로서, 각종 의미 관계를 바탕으로 다양한 추론기능을 제공하는 시맨틱 네트워크이다[4]. 여기에는 일상생활과 관련된 장소 관계, 물리적 관계, 사회적 관계, 시간적 관계, 심리적 관계 등으로 연결되어 있다. 본 논문에서는 ConceptNet에서 제공하는 추론함수와 특이성(기억에 남는 중요한 의미정보) 추론을 위한 베이지안 확률 모델의 분석을 통해 사용자 질의어와 라이프 로그 사이의 상관성을 분석하여 검색 결과를 제시한다.

첫 번째 단계는 확률 모델에 의한 연관 키워드 추출이다. 베이지안 네트워크 확률모델에서 키워드가 포함된 노드(x_i)와 직접적으로 연결된 노드를 찾고, 모델에 주어진 조건부 확률 파라미터(p_i)를 이용하여 다음과 같이 연관도(S_i)를 계산한다.

$$S_i = (p_i / 0.5) - 1.0$$

두 번째로, 사용자 로그로부터 사용자 질의와 관련이

높은 개념을 검색하기 위해서, ConceptNet Tool-kit [4]의 기능을 이용한 의미 기반 연관 키워드 추출이다. 이때 ConceptNet의 연관도 검색 함수(project_affectedive, project_consequences, project_details, project_spatial) 결과를 가중치 결합해서 사용하였다.

마지막으로 베이지안 네트워크와 ConceptNet을 이용해 추출된 키워드를 바탕으로 라이프 로그와 특이성을 검색한 다음, 키워드의 가중치(베이지안 네트워크에서 얻은 연결강도, ConceptNet에서 얻은 연관도)를 기준으로 선택된다.

3. 실험 및 결과

개발된 MyLifeBrowser의 의미 기반 검색 기능을 평가하기 위해 실제로 수집한 라이프 로그를 바탕으로 검색 실험을 수행하였다. 그림 10은 human1의 2월 25일 데이터에 대하여 '공부'라는 키워드로 단순 검색을 실시한 결과를 보여주며, 그림 11은 제안하는 확장 검색을 한 결과를 보여준다. 그림에서 검은색으로 표시된 점이 검색된 결과를 의미한다. 그림 10에서는 '공부'라는 키워드가 들어간 정보만 검색되기 때문에 적은 데이터만 검은색으로 표시된 것에 비해, 의미 기반 검색 결과에서는 도서관과 학교 내부에서의 일부 활동 궤적이 추가된 것을 알 수 있다.

표 1. '공부' 키워드에 관련된 BN 노드. 상위 3위까지

BN명	노드	관련노드	연결강도	순위
바쁨.BN	늦게까지공부	도서관	0.52	1
학교.BN	공부지역	도서관	0.5	1
음악.BN	공부중	대학교강의동	0.46	2
음악.BN	공부중	도서관	0.46	1
바쁨.BN	늦게까지공부	바쁨	0.36	3

표 2. 'study'를 컨셉넷으로 확장한 결과. 상위 3위까지

검색 키워드	확장 키워드	연관도
study	university	0.04
study	library	0.04
study	school	0.03

표 1은 베이지안 네트워크를 통해 '공부'에 관련된 특이성과 연결된 노드 가운데 상위 3위까지의 특이성을 나타낸다. 여기에서 연결 강도는 [-1,1]의 범위를 가진다. 표 2는 ConceptNet을 활용하여 'study'에 관련된 개념이 검색된 결과 중 일부를 나타낸다. 그 결과, 'university', 'library', 'school' 키워드가 추가된다. ConceptNet은 영어만 지원하기에 테이블 맵핑에 의해 영한 전환하였다.

4. 결론 및 향후 연구

본 논문에서는 모바일 디바이스로부터 수집한 사용자의 로그 정보를 의미 정보로 요약한 뒤, 이를 시각화하고, 추론용으로 개발된 베이지안 네트워크와 ConceptNet Toolkit을 활용한 의미 기반 검색 기능을 제안하였다. 사용자가 일상생활을 검색할 때 의미 기반 검색 기능이 유용할 것이다.

향후에는 검색 방법 및 결과에 대해 보다 넓은 도메인에서 체계적인 실험 및 평가가 필요할 것이다.

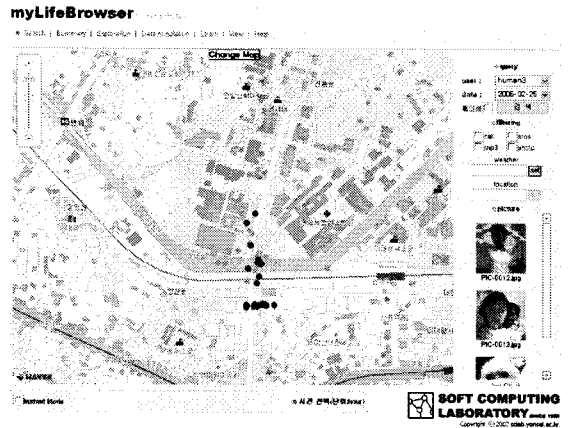


그림 10. '공부' 라는 키워드로 일반 검색 결과.

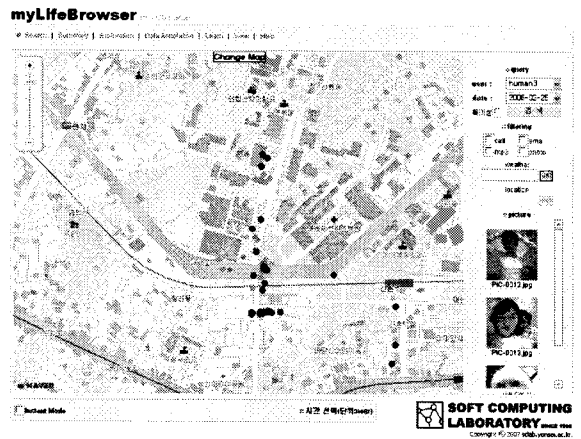


그림 11. '공부'라는 키워드로 의미 기반 검색 결과.

참고문헌

- [1] S. Ross, "Models of human memory," *Microsoft Research News & Highlights*, <http://research.microsoft.com/displayArticle.aspx?id=687>.
- [2] K.-J. Kim, Y.-S. Lee, K.-S. Hwang, J.-H. Hong, and S.-B. Cho, "Development of mobile life browser based on concept network," *Proc. of The 33st KISS Fall Conference*, vol. 33, no. 2, pp. 71-76, 2006.
- [3] K.-S. Hwang, S.-B. Cho, and J.-H. Lea, "A Bayesian inference model for landmarks detection on mobile devices," *Journal of Korea Information Science Society: Computing Practices*, vol. 13, no. 1, pp. 35-45, 2007. 02.
- [4] H. Liu and P. Singh, "ConceptNet - A practical commonsense reasoning tool-kit," *BT technology journal*, v.22 no.4, 2004, pp.211-226.