

소셜 네트워크 기반의 콘텐츠 추천 시스템의 개발

배운봉, 왕칭, 권경락, 손중수, 정인정
고려대학교 컴퓨터정보학과

e-mail : {beayunfeng, wangqing, helpnara, mis026, chung}@korea.ac.kr

Development of contents recommendation system based on social network

Yun-Feng Pei, Qing Wang, Kyung-Lag Kwon, Jong-Soo Sohn, In-Jeong Chung
Dept. of Computer and Information Science, Korea University

요 약

오늘날의 인터넷은 웹 2.0의 출현으로 인하여 콘텐츠의 생산주체가 서비스 제공자에서 서비스 수요자인 사용자들로 변화되고 있다. 이에 따라 사용자들의 경험은 콘텐츠의 품질에 큰 영향을 미치고 있으며 소셜 네트워크에서 취득한 콘텐츠는 검색으로 취득한 콘텐츠보다 신뢰를 받고 있다. 본 논문에서는 소셜 네트워크를 기반으로 사용자들에게 양질의 콘텐츠를 추천하기 위한 방법과 그 개발을 보인다. 소셜 네트워크는 XML 기반의 사용자 프로파일 기술 언어인 FOAF를 이용하여 수집하며 이를 통해 사용자와 사용자 사이의 관계를 수집한다. 그리고 웹 콘텐츠 출판언어인 RSS를 이용하여 각 사용자들이 블로그 등을 통해 배포한 콘텐츠를 수집한다. 본 논문에서 보이는 시스템은 FOAF와 RSS를 기초로 입력된 키워드에 대해 사용자와 콘텐츠의 관계를 분석하고 이를 통해 콘텐츠를 추천하는 기능을 가진다. 본 논문에서 보이는 시스템은 전통적인 콘텐츠 추천 시스템과 달리 사용자가 속한 소셜 네트워크에서 콘텐츠 생산자가 대한 중요도가 반영되므로 보다 신뢰성 있는 결과를 얻을 수 있다.

1. 서론

웹 2.0[1]은 새로운 유형의 웹 애플리케이션을 제공하는 플랫폼으로 지칭한다. 오늘날의 인터넷은 웹 2.0의 출현으로 인하여 콘텐츠의 생산주체가 서비스 제공자에서 서비스 수요자인 사용자들로 변화되었다. 이와 더불어 콘텐츠는 웹 2.0이 발전함에 있어서 가장 중요한 자리를 차지하게 되었다.

웹 2.0의 발전은 소셜 네트워크에 대하여 큰 영향을 미치고 있다. 소셜 네트워크는 개인을 의미하는 수많은 노드들이 구성된 사회적 관계망이다[2]. 오늘날의 소셜 네트워크는 데이터와 개인들의 네트워크를 결합하여 기존에 단순한 사회적 관계를 넘어 사용자들이 직접 참여할 수 있고 상호 작용이 가능하게끔 발전되고 있다[2]. 온라인 소셜 네트워크 서비스[3]는 웹 2.0 기술을 활용하여, 사용자들이 직접 콘텐츠를 작성하고 배포할 수 있어서, 이에 따라 사용자들의 경험은 콘텐츠의 품질에 큰 영향을 미치고 있으며 소셜 네트워크에서 취득한 콘텐츠는 검색으로 취득한 콘텐츠보다 신뢰를 받고 있다[4].

최근 웹 환경에서는 수많은 콘텐츠 사이트가 존재하고 방대한 콘텐츠 데이터는 기하적 수준으로 늘어나고 있다. 이에 사용자들을 하여금 원하는 콘텐츠를 검색하고 시간을 절약하며 사용자 만족도를 높이는데 대한 연구가 활발히 진행되고 있다. 전통적인 추천 방법에는 내용 기반 추천(Content-based Recommendation)

방법[5][9][10][11], 협업적 추천(Collaborative Recommendation)방법[6][12]들이 주로 사용되고 있다. 하지만 전통적인 추천방법들은 소셜 네트워크가 아닌 특정한 사이트나 도메인에만 적용되어 제한된 콘텐츠만을 사용자에게 제공한다.

본 논문에서는 사용자들에게 양질의 콘텐츠를 추천하기 위하여 FOAF[7]와 RSS[8]를 활용한 소셜 네트워크 기반의 콘텐츠 추천 시스템의 설계와 개발을 보인다. 소셜 네트워크를 XML 기반의 사용자 프로파일 기술 언어인 FOAF를 이용하여 수집하며 이를 통해 사용자와 사용자 사이의 관계를 수집한다. 그리고 웹 콘텐츠 출판언어인 RSS를 이용하여 각 사용자들이 블로그 등을 통해 배포한 콘텐츠를 수집한다. 본 논문에서 제안하는 시스템은 사용자가 입력한 키워드에 대해 사용자와 콘텐츠의 관계를 분석하고 이를 통해 콘텐츠를 추천하는 기능을 가진다.

본 논문에서 제안하는 시스템은 기존의 콘텐츠 추천 시스템과 달리 사용자가 속한 소셜 네트워크에서 콘텐츠 생산자에 대한 중요도가 반영되므로 보다 신뢰성 있는 결과를 얻을 수 있다.

본 논문은 2 장에서는 배경 기술과 관련 연구를 보이며 3 장에서는 소셜 네트워크 기반의 콘텐츠 추천 방법과 그 개발에 대해 보이며 4 장에서는 본 논문에서 보인 시스템의 적절성 및 효과 대해 논의한다.

2. 배경 기술 및 관련 연구

2.1. 배경 기술

FOAF(Friend-Of-A-Friend) 는 하나의 RDF 어휘집으로 XML 기반의 온톨로지 언어이다. 사용자의 정보를 기술한 FOAF 는 사람과 사람간의 관계, 사람과 개체들 간의 관계를 나타낸다. 이를 이용하여 소셜 네트워크 서비스(Social Network Service)에서 사용자들이 직접 수정한 개인 정보들이 FOAF 문서에 반영한다[7].

RSS(RDF Site Summary) 는 표준 웹 콘텐츠 출판언어이고 웹 상에서 주로 사용하는 콘텐츠 표현 방식이다. 기존의 전형적인 방식과 달리 RSS 를 이용한 사용자는 각각의 사이트 방문 없이 원하는 최신 콘텐츠들을 볼 수 있다[8].

2.2. 관련 연구

기존 추천 시스템에는 내용기반 추천방법을 이용한 시스템[5][9][10][11], 협업적 추천방법을 이용한 시스템[6][12]이 있다.

[9]에서는 각 노드에 포함된 아이템의 콘텐츠들의 미론적 관계 기반으로 유사도를 측정한다. 그래프에 포함된 노드들은 유사도를 이용하여 아이템에 대해 랭킹을 맺은 후 사용자에게 아이템을 추천한다. 하지만 소셜 네트워크가 아닌 그래프는 사용자의 참여도가 적어 이로 인해 제한된 아이템을 사용자에게 추천하는 단점이 있다.

[10]에서는 시맨틱 기술을 이용하여 아이템과 사용자 프로필에 대해 온톨로지를 구축한다. 구축된 온톨로지에 대하여 유사도를 측정한 후 사용자에게 추천한다. 하지만 실용적인 온톨로지를 구축하기 위해서는 도메인의 지식을 개념화할 수 있는 도메인 지식 전문가와 이를 구조화할 수 있는 온톨로지 설계 전문가가 필요하여[13] 이로 인해 다양한 아이템을 추천하지 못한 단점이 있다.

[11]에서는 아이템의 상세한 정보와 사용자 프로필의 관심사를 이용하여 콘텐츠의 다양한 아이템을 사용자에게 추천할 수 있으나 아이템에 대한 중요도를 고려하지 않았다, 그러나 소셜 네트워크는 아이템의 중요도에 대하여 쉽게 반영할 수 있다[14].

[12]에서는 사용자와 주변 사람의 프로필 유사도를 계산한 후 사용자들이 선호하는 태그를 이용하여 사용자에게 아이템을 추천한다. 하지만 소셜 네트워크는 아이템의 선호도에 대하여 보다 높은 영향을 준다 [15].

3. 본론

3.1. 소셜 네트워크 기반 콘텐츠 추천 기법

본 논문에서 제안하는 콘텐츠 추천 기법은 웹 상에 존재하는 사용자의 FOAF 및 RSS 를 수집하여 소셜 네트워크를 구축한다. 시스템은 이를 기반으로 사용자에게 의해 입력된 키워드를 통하여 소셜 네트워크 분석을 한 후 키워드에 관련된 콘텐츠를 추천한다.

소셜 네트워크 기반으로 콘텐츠를 추천하는 단계는 다음과 같다.

Step1) 시스템은 입력된 사용자의 아이디를 통해 FOAF 및 RSS 를 수집한 후 FOAF 의 <foaf:knows>를 통해 사용자와 관련된 사람들의 FOAF 및 RSS 를 수집한다.

Step2) 수집된 FOAF 및 RSS 를 이용하여 사용자와 관련된 사람들의 아이디, 사용자의 카테고리 및 콘텐츠등 필요한 정보를 추출한 후 사용자 중심으로 데이터베이스에 저장하면 콘텐츠가 포함된 소셜 네트워크가 생성된다.

Step3) 시스템은 사용자가 입력한 키워드를 통하여 소셜 네트워크 기반으로 관련된 사람의 FOAF 및 RSS 를 찾은 후 FOAF 문서 중 <foaf:interest>에 포함된 내용 및 RSS 문서 중 <category>에 포함된 내용과 비교한다. 만약에 입력한 키워드가 위의 두 태그 중에 모두 존재하면 시스템은 자동적으로 그 사람에게 가중치 2 를 부여하고, 하나만 존재하면 가중치 1 를 부여하며, 존재하지 않을 경우에는 가중치 0 을 부여한다.

Step4) 키워드에 관련된 콘텐츠를 Step2 에서 생성된 소셜 네트워크에서 추출한 후, 시스템은 추출된 콘텐츠에 대해 Step3 에서 계산된 가중치를 이용하여 콘텐츠 중요도를 계산한다. 본 논문에서 콘텐츠 중요도는 식(1)과 같이 계산된다.

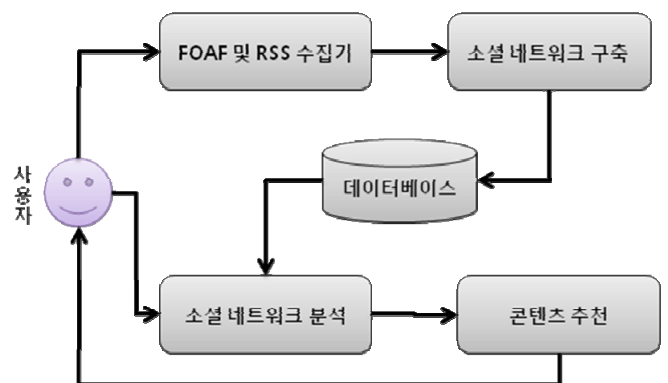
$$\text{콘텐츠 중요도} = \text{진입계수} \times \text{가중치} \quad (1)$$

식(1) 에서 진입계수는 소셜 네트워크에서 콘텐츠를 배포한 사람과 연결이있는 사람들의 수이다.

Step5) 시스템은 계산된 콘텐츠 중요도에 따라 추출된 콘텐츠 중에서 높은 값을 갖는 순서대로 추출하여 사용자에게 출력한다.

3.2. 시스템 아키텍처

본 논문에서 구축한 시스템의 아키텍처는 (그림 1) 과 같이 구성된다.



(그림 1) 시스템 아키텍처

FOAF 및 RSS 수집기는 사용자가 입력한 아이디를 통해 사용자의 FOAF 및 RSS 를 수집함과 동시에 사용자와 관련 있는 사람들의 FOAF 및 RSS 도 함께 수집한다. 소셜 네트워크 구축에서는 수집된 FOAF 및 RSS 를 통해 사용자를 중심으로 사용자에게 대한 소셜

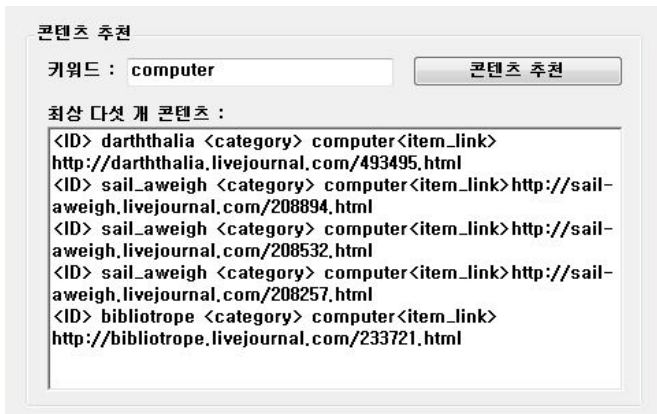
네트워크를 구축하고 데이터베이스의 각 테이블에 저장한다. 소셜 네트워크 분석에서는 사용자가 입력한 키워드를 통해 구축된 소셜 네트워크를 기반으로 소셜 네트워크를 분석한 후 키워드에 관한 콘텐츠를 데이터베이스에서 추출한다. 콘텐츠 추천에서는 추출된 매개 콘텐츠에 대하여 콘텐츠 중요도를 계산하여 최상위 다섯 개의 콘텐츠를 추출한 후 사용자에게 추천한다.

3.3. 시스템 개발

제안한 시스템의 개발환경은 마이크로소프트 비주얼 스튜디오 2008 과 마이크로소프트 SQL 서버를 이용하였으며 개발언어는 C#를 이용하였다.



(그림 3) FOAF 및 RSS 수집과 소셜 네트워크 구축 인터페이스



(그림 4) 콘텐츠 추천 인터페이스

제안한 시스템의 FOAF 및 RSS 수집과 소셜 네트워크 구축은 (그림 3)과 같이 사용자가 입력한 아이디를 이용하여 FOAF 와 RSS 를 수집한 후 사용자에게 관한 소셜 네트워크를 구축한다. 사용자의 참여도가 많은 소셜 네트워크를 이용하면 제한된 콘텐츠만을 추천하는 기존 그래프 기반 콘텐츠 추천 방법과 달리 다양하고 범용적인 콘텐츠를 사용자에게 추천할 수 있다.

(그림 4)에서 보이는 콘텐츠 추천 인터페이스는 시스템이 사용자가 입력한 키워드를 바탕으로 3.1 절에서 제안한 기법을 이용하여 구성된 소셜 네트워크를

분석한 후, 사용자에게 키워드에 관련된 콘텐츠 중에서 높은 값을 갖는 순서대로 다섯 개의 콘텐츠를 추천한다.

(그림 3)은 사용자가 FOAF 및 RSS 수집과 소셜 네트워크 구축을 이용하여 사용자의 아이디 mieza 를 시스템에 입력하고 시스템은 사용자 중심으로 소셜 네트워크를 구축하는 모습을 나타낸다. 구축된 소셜 네트워크에는 사용자와 관련 사람들의 아이디, 카테고리 및 콘텐츠 등의 정보들을 포함한다. 사용자는 콘텐츠 추천을 이용하여 찾고자 하는 콘텐츠에 대한 키워드 computer 를 시스템에 입력한다(그림 4). 시스템은 입력된 키워드를 이용하여 구축된 소셜 네트워크를 바탕으로 키워드에 관련된 콘텐츠를 추출하고 이에 대하여 콘텐츠 중요도를 계산한 후, 높은 값을 갖는 순서대로 콘텐츠를 정렬한다. 본 시스템에서는 높은 값을 갖는 다섯 개의 콘텐츠를 사용자에게 추천해준다. 추천된 매개 콘텐츠 중에는 콘텐츠를 작성한 사람의 아이디, 콘텐츠가 포함된 카테고리, 콘텐츠에 대한 링크가 포함된다. 사용자는 추천된 콘텐츠의 링크를 이용하여 보다 쉽게 콘텐츠를 접근할 수 있다.

본 논문에서 구현된 시스템은 콘텐츠에 대하여 콘텐츠 중요도를 계산하였기 때문에 기존 콘텐츠 추천 방법과 달리 콘텐츠에 대한 중요도를 고려하여 사람들이 선호하고 보다 신뢰성이 높은 콘텐츠를 사용자에게 추천할 수 있다.

4. 결론

본 논문에서 구축된 시스템은 소셜 네트워크를 바탕으로 사용자의 FOAF 및 RSS 를 수집하고, 3.1 절에서 제안한 소셜 네트워크 기반 콘텐츠 추천 기법을 통해 입력된 키워드에 관련된 콘텐츠를 추천하는 시스템을 구현하였다. 오늘날의 콘텐츠의 생산주체가 서비스 제공자에서 서비스 수요자인 사용자들로 변화되었다. 사용자들의 경험은 콘텐츠의 품질에 큰 영향을 미치고 있을 뿐만 아니라, 소셜 네트워크에서 획득한 콘텐츠는 검색을 통해 얻은 콘텐츠보다 상대적으로 높은 신뢰도를 가지고 있다. 따라서 본 논문에서는 입력된 키워드와 관련된 콘텐츠에 대하여 콘텐츠 중요도를 계산한 후 상위 콘텐츠를 분석하는 시스템을 제안하고 구현하였다.

본 논문에서 제안한 방법을 활용하면 신뢰성이 있는 콘텐츠를 사용자에게 추천할 수 있을 뿐만 아니라 소셜 네트워크 바탕으로 사용자들간의 콘텐츠 공유를 더욱 활발하게 촉진 시킬 수 있다.

참고문헌

[1]O'REILLY. "What is Web 2.0"
 [2]AL Barabási, H Jeong, Z Néda, E Ravasz, A. Schubert, T. Vicsek. "Evolution of the social network of scientific collaborations"
 [3]www.livejournal.com

- [4]R Kumar, P Ragbavan, S Rajagopalan, A Tomkins. “The web and social networks”
- [5]M. Claypool, A. Gokhale, T. Miranda, P. Murnikov, D. Netes and M. Sarti. “Combining Content-Based and Collaborative Filters in an Online Newspaper”
- [6]Conor Hayes, Pádraig Cunningham. “Context boosting collaborative recommendations ”
- [7]<http://www.foaf-project.org>
- [8]RSS Specifications.
<http://www.rssboard.org/rss-specification>
- [9]Tao Peng, Wendong Wang, XiangYang Gong, Ye Tian, XiaoGang Yang, Jian Ma. “A Graph Indexing Approach for Content-Based Recommendation System ”
- [10]Gao, Yuhong Li, Li Han, Jian Ma. “InfoSlim: An Ontology-Content Based Personalized Mobile News Recommendation System ”
- [11]Michael J. Pazzani, Daniel Billsus. “Content-based Recommendation Systems”
- [12]Huizhi Liang, Yue Xu, Yuefeng Li, Nayak R. “Collaborative Filtering Recommender Systems Using Tag Information”
- [13]조이현, 박대원, 박동훈, 문홍구, 권혁철. “비전문가에 의한 상하위 관계 중심의 온톨로지 공동구축 방법”
- [14]Ming-Hsiung Ying, Hao-Hsuan Huang. “Item Selection Strategic via Social Network Analyze”
- [15] Claudia Canali, Michele Colajanni, Riccardo Lancellotti. “Characteristics and evolution of content popularity and user relations in social networks”