

분산처리 환경에서 소셜 네트워크를 이용한 전문가 추천 시스템

한승민, Mohamad Mehidi Hassan, 허의남
경희대학교 컴퓨터공학과
e-mail: {han905, hassan ,johnhuh}@khu.ac.kr

An Export Recommendation System using Social Network in Distributed Processing Environment

Seung-Min Han, Mohamad Mehidi Hassan, Eui-Nam Huh
Dept of Computer Engineering, Kyung Hee University

요 약

최근 소셜 네트워크 기반의 서비스들을 중심으로 거대한 시장이 형성되고 있다. 소셜 네트워크 서비스의 수가 급격하게 증가하고 있는 상황에서 소셜 네트워크에 존재하는 수많은 리소스들을 이용하여 여러 영역에서 필요한 전문가를 찾아주는 시스템을 구성해보고자 한다. 기존의 추천시스템은 특정한 상품이나 소수 사용자 피드백을 바탕으로 추천 시스템을 구성하였다. 그러나 소셜 네트워크를 이용한 추천 시스템에 관한 연구는 미비한 상황이다. 본 논문에서는 거대한 소셜 네트워크의 리소스를 처리하기 위한 분산처리시스템과 소셜 네트워크를 이용하여 전문가 추천 시스템을 구성한다. 이를 이용하여 다양한 영역에서 적절한 전문가를 찾아주는 서비스를 제공해 준다.

1. 서론

최근 들어, 사회는 좀 더 복잡해지고 다양해지며 매우 빠른 변화에 적응해가고 있다. 이런 사회의 구성원들은 좀 더 자신의 가치를 유지하기 위해 타인들과 많은 관계를 유지하길 원하는 경향이 있는데 자신을 표현하고 정체성을 유지하기 위한 기본적인 욕구가 IT기술과 접목되어 FaceBook, MySapace, Cyworld와 같은 SNS(Social Networking Service)를 탄생하게 하였고 SNS 시장을 급격히 성장하게 하였다. 실제 미국 SNS 업계 1위인 MySpace 회원 수는 2억 명을 넘어섰고 최근 급성장 중인 Facebook의 경우 매년 방문자 수의 증가율이 200%에 달한다. SNS 기업을 비롯해 주요 IT기업들은 잠재적 소비자들이 가득한 SNS 시장을 각종 수익모델과의 연계로 나서고 있어 새로운 광고 및 콘텐츠, 서비스를 제공해주는 통로로써 중요성이 매우 높아지고 있다.[1]

소셜 네트워크(Social Network)이란 사람과 사람 사이의 연결망으로 여러 가능한 관계를 의미한다. Social Networking은 이런 관계를 만드는 행위를 의미하고 사람들 간의 관계 형성을 목적으로 하는 SNS는 Social Networking을 보다 쉽게 할 수 있도록 도와주는 역할을 한다. 관계 형성을 위한 SNS는 현재 수백 개에 이르며 각 분산되어 있는 소셜 네트워크를 이용하여 의미 있는 서비스를 제공하려 한다. 수많은 소셜 네트워크에서 적절한 리소스를 찾기 위해서는 분산처리 환경은 필수[1]이다. 거대한 인맥 네트워크들의 내용을 분석하고 분류해내

기 위해 본 논문에서는 분산처리 환경을 기반으로 한다.

현실 세계에서 사람들은 어떤 사건이 생기면 자신의 인맥을 이용하여 그 사건을 해결하기 위한 해결책 또는 조언을 받길 원한다. 이런 경향을 이용하여 질의응답기반의 지식 서비스들이 큰 인기를 얻고 있는 상황이다. 이런 서비스들을 보면 어떤 해결책이나 조언을 추천해주는 기능을 가진다. 기존에 이러한 추천 시스템은 고객이 만족할 만한 상품(생산품, 영화, 음악 등)들을 추천해주는 시스템에 관한 연구만 중점적으로 진행되고 있다.

따라서 본 논문에서는 기존의 추천시스템을 살펴보고 소셜 네트워크를 이용하여 사용자가 필요로 하는 분야의 전문가를 추천해주는 시스템을 제안해 보고자 한다. 이러한 추천 시스템은 새로운 서비스의 창출로 이어지고 새로운 서비스의 탄생은 수많은 비즈니스 모델 창출에도 기여할 수 있다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 추천 시스템에 대해 설명하고, 3장에서는 소셜 네트워크를 이용한 전문가 시스템을 구성해보았다. 4장에서는 결론과 향후 과제에 대해 논의한다.

2. 관련 연구

2.1 추천 시스템을 위한 기술

(1) Demographic Recommender Systems[3]

기본적으로 사용자들의 속성(attribute)을 이용하며 사

용자 통계학적인 자료를 바탕으로 등급을 나누어 추천을 해주는 방식이다. 즉 인구통계적인 계산을 바탕으로 사용자들의 속성 분석하여 필요한 정보를 제공해주게 된다.

사용자의 속성만을 가지고 추천을 해주므로 사용자가 활동한 수많은 로그를 분석하지 않아도 되어 시스템에 적용이 비교적 용이하나 개개인의 분류(classification)에 대한 어려움이 있고 통계학적인 자료를 바로 바로 업데이트 하고 분류하여 반영하는 어려움으로 인하여 추천 결과가 부정확한 경우가 발생하게 된다.

(2) Contents-based Recommender Systems[3][5]

콘텐츠 기반의 필터링 기술은 속성이 비슷한 상품을 분류하여 사용자가 원하고자 하는 상품을 추천해주는 기술이다. 사용자들은 상품과 연관된 상품에 대한 추천 정보를 제공 받을 수 있지만 아래와 같은 문제가 있다.

- Over-specialization : 속성이 비슷한 상품을 추천해주는 데 있어 중복된 상품에 대한 정보를 제공받는 경우가 있다.
- Limited content analysis : 사용자가 만족하는 상품의 정보를 충분히 제공하지는 못한다.
- New user problem : 새로운 사용자는 기존의 프로파일 정보의 부재로 원하지 않는 추천을 받게 된다.

(3) Collaborative Recommender Systems[3][4][5]

협업 기반의 필터링을 이용한 추천 시스템은 대부분의 추천 시스템에서 사용하는 기술이다. 이는 상품에 대한 사용자의 의견이 바로 반영되어 비교적 적절한 추천을 해주기 때문이다. 비슷한 성향(취미, 흥미, 관심 등)을 가지고 있는 사용자들로 분류 시킨 후 기존 사용자들의 평가들을 바탕으로 현재 사용자에게 원하는 정보를 추천하게 된다.

하지만 “Cold-Start”[4], “Data Sparseness”[6] 라는 문제점을 가지고 있다. “Cold-Start” 는 새로운 상품이 들어오는 경우 기존의 유사한 속성을 지닌 상품보다 높은 품질을 지니고 있지만 추천 시스템은 기존의 반영된 사용자의 의견을 기반으로 추천해주므로 새로운 상품에 대한 관심은 떨어지게 된다. “Data Sparseness”는 너무 많은 상품에 대한 평가로 인하여 발생하는 문제이다. 상품들에 대한 평가는 종종 바뀌고 사용자의 수는 상대적으로 적어서 추천받고자 하는 상품과 비슷한 상품을 찾는 사용자들에게 혼란을 준다.

(4) Hybrid Recommender Systems[4][5]

하이브리드 추천 시스템은 위에서 언급한 콘텐츠 기반(Contents-based)과 협업(collaborative)기반의 필터링 기술을 함께 취한 것으로 두 가지 기술을 모두 취합하거나 한 가지 방법을 사용하여 나온 결과를 다른 기술을 사용하여 처리하거나 두 가지 기술의 특징만을 취하거나 하는 형태로 여러 가지 다양한 방법들이 존재하며 여전히 연구 진행 중이라고 할 수 있다.

2.2 Social Network를 이용한 시스템

최근 소셜 네트워크에 대한 연구는 이메일(E-mail)[8], 정보추천(information-recommendation)[9], 지리(geography)[9] 혹은 웹 서비스(Web-service)[7] 등 여러 분야에 적용되고 있다.

P. Oscar Boykin이 제안한 시스템[8]을 보면 최근 급격히 증가하는 스팸메일로 인하여 불필요한 리소스의 낭비뿐만 아니라 메일에 대한 신뢰도 역시 떨어지고 있는 문제를 소셜 네트워크를 사용하여 해결하는데 사용자의 메일함의 메일 정보를 바탕으로 망을 구성하여 신뢰도를 평가하여 스팸메일을 걸러주는 시스템이다. 또한 Quentin Jones가 제안한 P3 시스템[9]은 지리정보를 바탕으로 같은 지역의 소셜 네트워크를 구축하고 이를 이용하여 지역 내의 사용자들에게 지리기반의 이점을 이용한 서비스를 제공해주게 된다. 이처럼 다양한 방면을 사용하는 소셜 네트워크를 이용하여 도움이 필요한 영역의 전문가 추천 시스템을 구축하여 연구 발전을 해보고자 한다.

3. 전문가 추천 시스템 설계

본 논문에서는 소셜 네트워크를 이용하여 각 사용자가 원하는 상황에서 적절한 전문가를 추천해주는 시스템을 제안하려 한다. 이를 기반으로 의료분야에서는 환자를 위한 의사 추천 시스템, 특별한 교육이 필요한 학생에게는 선생님 추천 시스템, 컴퓨터 수리가 필요한 사용자에게는 컴퓨터 수리 전문가 추천 시스템 등 다양한 방면에 활용하여 다양한 비즈니스 모델의 창출에 기여한다.

3.1 R-Rank(Recommendation Rank) 구축

소셜 네트워크는 정점(vertex)와 점간의 관계로 이루어져 있다. 이 가운데 소셜 네트워크를 구성하는 가장 중요한 것은 관계이다. 이 관계는 행위(activity)에 의하여 이루어지며 우리는 관계를 이용하여 R-Rank 구축하여 사용자가 필요로 하는 적절한 전문가를 추천해 주며 다음(수식 1)을 이용하여 추출해 낼 수 있다.

$$R-Rank = (rl + \frac{Sreq}{tot}) \times PD$$

(수식 1) R-Rank

기본적으로 소셜 네트워크를 이용한 전문가 추천 시스템은 자신을 기준으로 검색을 해나가게 된다. (수식 1)에서 rl 은 네트워크 라인의 개수이다. 즉 사용자의 인맥의 개수를 의미한다. $Sreq$ 는 기존에 추천시스템을 통해 들어온 요청을 해결해준 횟수이며 tot 요청이 들어온 총 횟수이다. 또한 중요한 가중치로 프로필 데이터를 통해 제공된 전문성을 알려주는 상수로 PD 가 존재하며 추천하려

는 특정 분야의 사용자 프로필 데이터를 바탕으로 PD는 (수식 2)와 같다.

$$PD = match(userquery) \times \frac{2^{depth}}{10}$$

match(userquery): Search value about user query
depth: Relation depth value

(수식 2) PD

소셜 네트워크에 존재하고 있는 사용자는 모두 Profile 을 지닌다. 이런 Profile을 분석하여 전문가인지 아닌지를 판단하게 된다. 시스템에 의해 추천을 받은 전문가는 전체 요청 중에 승낙한 요청 비율을 계산한다. 계산된 값은 사용자가 추천을 원하는 질의에 따라 그 사람의 특정 분야에 전문성이 있는지 판단하며 자신과 관계가 가까울수록 신뢰성이 있다고 판단하여 좀 더 높은 값을 지니게 된다. 만약 자신이 속한 소셜 네트워크에서 원하는 전문가가 없다면 PD 값은 0이 되어 추천 검색 시 제외되며 매칭이 된다면 R-Rank를 구해 시스템에 의해 사용자에게 노출 될 수 있다.

3.2. 정보 추출을 위한 소셜 네트워크 순환 알고리즘

전문가 추천 시스템에 핵심 가중치인 R-Rank를 구축하기 위해서 기존의 소셜 네트워크를 검색하면서 관련된 정보를 추출하여야 한다. (그림3)은 스택을 이용한 소셜 네트워크 순환 알고리즘 이다.

```

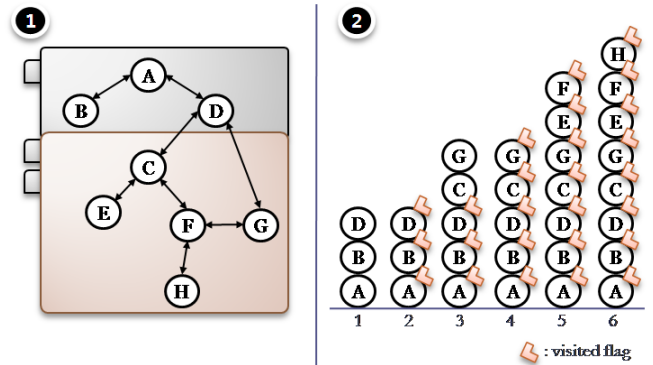
Procedure Traverse SN(userVertex)
initial : Stack S[0-N]
Begin
S.push(userVertex)
userVertex.visited = True

For(i=0; i<userVertex.neighborSize; i++)
Begin
userVertex.neighbor[i].visited = True
S.push(userVertex.neighbor[i])
End

For(j=1; j<S.stackSize; j++)
Begin
currentVertex = S[j]
If currentVertex equal True then
continue loop
Else
For(k=0; k<currentVertex.neighborSize; k++)
Begin
If currentVertex.neighbor[k] equal False then
S.push(currentVertex.neighbor[k])
currentVertex.neighbor[k].visited = True
EndIf
End
EndIf
End
End
    
```

(그림 1) 소셜 네트워크 순환 알고리즘

위 알고리즘에서 우선 시작은 자신이 된다. 내가 속한 소셜 네트워크들을 이용하여 검색을 하게 되는데 처음 나와 내 이웃들이 스택에 추가되면 내 정보는 스택 가장 아래 위치하게 되고 나와 강한 인맥을 형성할수록 나와 가까워지게 된다. 스택에 쌓인 정보를 가지고 전문가를 추천할 수가 있으며 아래 그림의 과정과 같다.

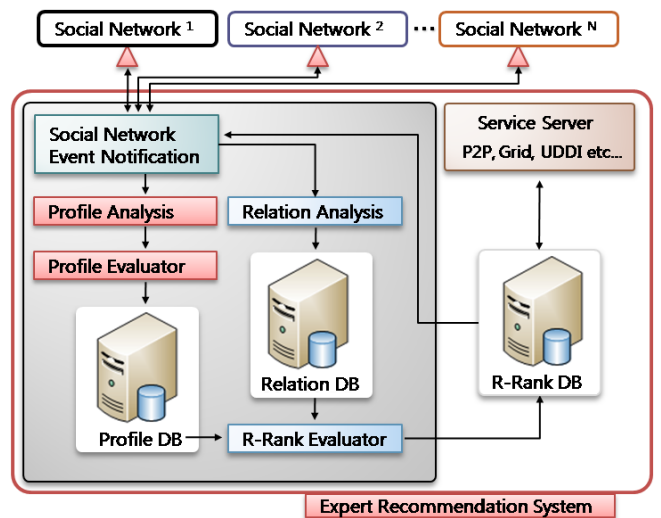


(그림 2) 소셜 네트워크 순환 알고리즘 예제

(그림 2) 1번 단추부분의 그래프 구조를 예를 들어 보면 1번 그래프에서 한 개의 태그로 된 박스 먼저 순회를 시작한다. 만들어진 자료를 바탕으로 태그 두 개로 이루어진 박스를 순회하여 2번과 같은 구조로 자료를 스택에 넣고 R-Rank를 추출해 낸다.

3.3 전문가 추천 시스템 구조

아래 (그림 3)은 분산처리 환경에서 특정 분야의 전문가 서비스 제공을 위한 전문가 추천 시스템의 구조를 나타낸 것이다. 각 소셜 네트워크에서 새로운 관계가 발생하면 Event Notification은 이벤트를 받아 처리하게 된다.



(그림 3) 전문가 추천 시스템

다음은 (그림 3) 시스템에 블록들에 대한 설명이다.

• Social Network Event Notification

소셜 네트워크에서 새로운 이벤트를 감지하여 기존의 랭킹을 변화시켜주거나 사용자의 질의를 받아 R-Rank DB를 이용할 시에 원하는 정보가 없으면 요청한다.

• Profile Analysis

사용자의 질의를 바탕으로 Profile내에 적합한 내용이 있는지 분석한다. 확장하여 사용자가 속해있는 네트워크에서 IR(Information Retrieval)을 통해 정확성을 높일 수 있다.

• Profile Evaluator

Profile analysis에서 얻은 정보를 평가하는 부분으로 평가 방법을 시스템에 맞게 세분화하여 정확성을 높일 수 있다.

• Relation Analysis

사용자가 속해있는 여러 소셜 네트워크마다 존재하는 사용자의 관계를 분석한다.

• Relation Evaluator

분석된 관계를 통해 전문가 요구 요청이 들어온 것과 사용자의 만족도를 평가한다.

• R-Rank Evaluator

분석 평가된 관계와 평가된 사용자 Profile 데이터를 바탕으로 R-Rank를 계산한다.

• Service Server

사용자의 질의가 요청되면 서비스 서버는 R-Rank DB에 질의를 요청한다. R-Rank DB는 기존에 검색되어진 데이터가 있는지 찾아보고 없다면 Social Network Event Notification 모듈에 요청해준다.

3.4 전문가 추천 시스템 평가

기존의 추천 시스템들을 통해서도 사용자가 원하는 특정 영역에서의 전문가들을 검색해 줄 수가 있다. 하지만 이를 위해서는 그 전문가에 대한 기반 데이터가 존재하여야 하고 이를 평가한 항목들이 있어야 하는 큰 단점이 있지만 본 논문에서 제안한 시스템은 기반 데이터를 미리 구축하지 않아도 자신이 속한 소셜 네트워크를 통해 전문가를 추천해 줄 수가 있다.

본 논문에서는 기존의 Contents-based, Collaborative 필터링 기술을 혼합한 Hybrid 필터링 기술을 사용한 추천 시스템의 개념을 확장하였고 기존 시스템의 문제점을 다음 <표 3>과 같이 해결함을 알 수 있다.

<표 3> 소셜 네트워크 이용 시 기존 문제점에 대한 해결책

문제점 \ 기술	Using Social Network
Cold-Start	소셜 네트워크 참여로 해결
New-member	소셜 네트워크 의 프로파일 정보 이용
Over-specialization	사용자 프로파일에서 분류가 되어 있음
Data Sparseness	R-Rank 이용

4. 결론

본 논문에서는 기존의 추천 시스템들을 알아보고 여러 문제점을 해결한 Hybrid 추천 시스템에 소셜 네트워크를 도입하여 기존의 Hybrid 추천 시스템보다 신뢰도와 정확성을 높이는 것을 목표로 하였다.

전문가 추천 시스템은 사용자가 직면한 환경에 따라서 다양한 속성을 통해서 발전 가능한 응용 시스템이다. 기존의 소셜 네트워크를 이용하여 상품을 추천해주는 연구들은 진행이 되었지만 어떤 분야의 전문가를 추천해주 것은 새롭게 시도되는 연구라는 점에서 평가가치가 있다. 향후 과제로 R-Rank를 좀 더 발전 시켜 정확성을 높이는 연구와 추천 시스템을 비즈니스 모델과 결합하여 더 많은 서비스를 위한 바탕이 될 수 있도록 하는 연구가 기대된다.

Acknowledgement

"본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT 연구센터 지원사업의 연구결과로 수행되었음" (IITA-2008-C1090-0801-0040)

참고문헌

[1] "SNS의 변화하는 서비스 전략", 한국소프트웨어진흥원, 2007

[2] M. Lan and W. Zhou "An Adaptive Information Grid Architecture for Recommendation System", School of Information Technology, Deakin University, 2003

[3] Iváan Cantador¹, Miriam Fernández¹ and Pablo Castells¹, "A Collaborative Recommendation Framework for Ontology Evaluation and Reuse", Universidad Autónoma de Madrid, Spain, 2006

[4] J. Ben Schafer, Dan Frankowski, Jon Herlocker and Shilad Sen, "Collaborative Filtering Recommender Systems", The Adaptive Web, 2007

[5] Gediminas Adomavicius and Alexander Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions", IEEE Educational Activities Department, 2005

[6] Jennifer Golbeck, "Computing with Trust: Definition, Properties, and Algorithms", Securecomm and Workshops, 2006

[7] 신동욱, 강진범, 양재영, 최중민, "사회망을 이용한 이종간 협업분야 탐지 시스템", 한국정보과학회, 2007

[8] P. Oscar Boykin and Vwani P. Roychowdhury, "Leveraging social networks to fight spam" IEEE Computer, 2005.

[9] Quentin Jones and Sureshini A. Grandhi, "P3 systems: Putting the place back into social networks" IEEE Internet Computing, 2005.