

제품관여 수준에 따라 소셜 정보가 추천 성능에 미치는 영향

송희석* · 주석정** · 이재훈***

The Effects of Social Information on Recommendation Performance According to the Product Involvement Level

Hee Seok Song* · Seok Jeong Joo** · Jae Hoon Lee***

Abstract

With the rapid increase of social network usage, there are emerging trends of adopting social information among online users in building recommendation system. This study aims to investigate whether the additional usage of social information can improve recommendation performance in recommendation system and how much the improvement can be different according to the product involvement level. As an experiment result, social information does not affect positively to the recommendation accuracy but affect significantly to the recommendation quality. Also social information contributed more sensitively to the improvement of recommendation quality in high product involvement domain.

Keywords : Social Information, Product Involvement Level, Recommendation, Recommendation Quality, Big Data, Mahout

논문접수일 : 2014년 11월 07일 논문수정일 : 2014년 11월 26일 논문게재확정일 : 2014년 11월 26일

※ 이 논문은 2014학년도 한남대학교 학술연구조성비 지원에 의하여 연구되었음.

* 한남대학교 경영정보학과 교수, e-mail : hssong@hnu.kr

** 교신저자, 한남대학교 경영정보학과 교수, e-mail : sjjoo007@nate.com

*** 창업진흥원, e-mail : edwardlee@kised.or.kr

1. 서 론

스마트기기의 급속한 보급과 클라우드 컴퓨팅 및 사물인터넷의 성장으로 촉발된 방대한 양의 구조화된 정형데이터와 소셜스트림 및 센싱데이터를 포함하는 비정형데이터는 고객과 환경을 이해하는 수단으로 활용되어 새로운 비즈니스 기회를 제공하고 있다. 이와 동시에 빅데이터는 정보과부하(Information overload)를 유발하여 사용자가 원하는 정보를 획득하기까지 많은 시간과 노력을 소비하게 한다. 이에 사용자의 선호 정보를 파악하여 맞춤형 정보를 추천해 주는 기능이 갈수록 중요시 되고 있다. 사용자의 선호도를 기반으로 상품이나 콘텐츠의 추천에 가장 많이 활용되고 있는 기법은 협업필터링(Collaborative Filtering) 기법이다. 협업필터링 기반의 추천은 구전을 통한 상품이나 서비스의 확산에 착안하여 개발된 기법으로써, 사용자에게 일부 콘텐츠를 제공하고 이에 대한 선호 점수를 입력받아 선호 패턴이 유사한 다른 사용자들이 선호했던 콘텐츠를 추천하는 방식이다[Sarwar, 2000]. 협업필터링 기법을 추천에 활용한 주요 연구로는 미네소타 대학의 GroupLens, 버클리 대학의 Jester, MIT의 음악 추천시스템 RINDO 등이 있으며, Net Perceptions사의 GroupLens Recommendation Engine, Gustos사의 Gustos Prediction Server 등 다양한 상용시스템이 존재한다.

최근 소셜 네트워크 사용이 보편화되면서 소셜 정보를 추천에 활용하는 시도가 급증하고 있다. 여기서 소셜 정보란 소셜 네트워크상의 사용자 간 관계정보를 총칭하는 것으로 사용자 간 명시적인 친구관계, 신뢰관계, 또는 사용자간 상호작용에서 유도된 암묵적인 관계 등 다양한 형태의 사용자 간 관계형태를 의미한다. 소셜 정보를 추천에 활용하면 친구 또는 나와 특별한 관계를 맺고 있는 사람의 글이나 구매제품을 추천할 수 있기 때문에 추천 이유를 설명하기 용이하여 추천

결과의 신뢰도가 높아질 수 있다. 소셜 네트워크의 활성화와 더불어 웹 2.0의 확산에 따른 정보의 개방 및 공유 트렌드로 인해 사용자 간 소셜 정보의 수집이 용이해지면서 추천시스템 연구 분야에서도 다양한 소셜 정보를 추천시스템에 활용하여 추천정확도와 추천품질을 개선하고자 하는 시도가 증가하고 있다. 이 중 다수의 연구에서는 소셜 정보의 추가적인 활용이 추천 성능 향상에 기여하고 있는 것으로 보고하고 있으나, 다른 일부 연구에서는 소셜 정보의 추가적인 활용이 추천품질 향상에 의미 있는 수준으로 기여하지 못하는 것으로 보고하는 등 소셜 정보의 추가적인 활용이 추천 성능에 미치는 영향에 대해서는 상반된 의견이 존재한다. Arazy et al.[2009], Ricci et al.[2011]과 Neto and Nowe[2012]의 연구에서도 소셜 정보를 활용한 추천시스템이 주목받는 연구토픽이라고 지적하며, 많은 연구자들이 소셜 추천시스템과 기존 추천시스템에 대한 추천품질 비교결과를 제시하고 있지만 아직 연구자들 사이에서 소셜 추천시스템의 성능이 더 좋다는 사실에 대해서는 공감대 형성되어 있지 않은 상황이라고 언급하고 있다.

추천시스템에 대한 성능 비교연구 결과가 연구마다 다르게 나타나는 것은 어떤 도메인에서 어떤 소셜 정보를 어떤 방법으로 추천에 활용하는가의 차이 때문에 발생한다. 대부분의 추천시스템 연구들은 저관여제품(영화, 음악 등) 또는 고관여제품(가구, 전자제품 등)과 같은 특정 추천도메인을 정하여 추천시스템을 설계하기 때문에 연구자마다 서로 다른 도메인에서 얻어진 실험결과를 정리하여 소셜 정보 활용의 효과를 제시함으로써 소셜 추천시스템의 성능에 대한 엇갈린 결과가 나타날 수 있다. 또한 연구자들마다 소셜 정보를 친구관계, 신뢰관계, 명성 등 다양한 형태로 정의하고, 서로 다른 형태의 소셜 정보를 추천에 활용하였기 때문에 소셜 정보의 활용효

과가 다르게 나타났을 가능성도 존재한다. 이외에도 소셜 정보를 추천에 활용하는 방식도 연구자마다 상이하여 소셜 추천시스템의 성능평가 결과가 달라졌을 가능성도 있다. 따라서 본 연구에서는 다양한 소셜 정보를 동일한 조건 하에서 실험하여 소셜 정보의 추가적 활용이 추천시스템 성능개선에 효과적인지를 밝히고자 한다. 보다 구체적으로 본 연구의 목표는 다음과 같다. 첫째, 신뢰, 친밀성, 관계지속성, 전문성 등 다양한 형태의 소셜 정보의 추가적인 활용이 추천 성능 향상에 과연 기여하는지를 살펴보고 둘째, 소셜 정보의 추가적인 활용이 추천 성능에 미치는 영향은 제품관여 수준에 따라 어떻게 달라질 것인지를 살펴보는 것을 본 연구의 목표로 한다. 대부분의 기존 연구는 몇몇 소셜 정보의 추가적인 활용이 추천시스템 성능 향상에 도움이 되는지를 확인하는 단편적인 연구가 대부분이며, 도메인 별로 다양한 소셜 정보의 추가적인 활용이 추천 품질 향상에 얼마나 기여하는지를 종합적으로 살펴보는 연구는 찾아보기 어렵다. 본 연구의 구성은 다음과 같다. 먼저 제 2장에서 두 가지 핵심 연구문제를 제시하고 제 3장에서 용어정의를 통해 연구 문제를 정의하며 제 4장에서는 실험에 필요한 소셜 추천시스템을 제안하며, 제 5장에서는 실험결과를 제시하며, 마지막으로 제 6장에서 결론을 맺는다.

2. 연구문제

본 연구에서는 다음과 같은 두 가지 연구문제를 제시하고 실증연구를 수행하는 것을 목표로 한다. 첫 번째 연구문제는 신뢰, 친밀성, 관계지속성, 전문성 등 다양한 형태의 소셜 정보의 추가적인 활용이 추천 성능 향상에 과연 기여하는지를 살펴보는 것이다. 응용심리학과 마케팅 연구분야의 행위이론(Behavioral theory)에 의하면

영향자(Influencer)와 수용자(Recipient) 사이의 상호작용의 빈도, 관계기간, 영향자의 명성과 영향자에 대한 신뢰 등 다양한 형태의 사용자 관계에 따라 수용자의 의사결정이 달라질 수 있음을 주장하고 있다. Gilly et al.[1998]은 영향자와 수용자의 인지적 유사성이, Levin과 Cross[2004]는 상호작용 빈도와 관계기간 등의 관계 강도가, Smith et al.[2005]은 수용자의 영향자에 대한 신뢰가 각각 수용자의 의사결정에 영향을 준다고 하였다. Bonhard[2005]는 기존 연구에 대한 질적 연구를 수행하여 추천인과 피추천인 간 관계가 의사결정에 중요한 영향을 미치는 것을 확인하기도 하였다. Groh and Ehmig[2007]에 의하면 일반 사용자의 의사결정은 지인의 충고를 기반으로 하는 경우가 많으며, 특히 기호와 관련된 도메인의 경우 사회적 관계가 의사결정에 중요한 영향을 미친다고 하였다. 이러한 주장은 기계학습 분야의 공학연구에서 실험을 통해 확인되기도 했는데, 기존연구에 따르면 소셜 정보의 추가적인 활용은 평가데이터의 희박성 문제를 완화하는데 도움을 줌으로써 추천 정확도를 개선할 수 있다는 점에서 소셜 정보의 활용이 추천시스템 정확도 개선에 효과적이라는 것이다[Kim et al., 2011; Zheng and Li, 2011; Carrer-Neto et al., 2012]. Neto and Nowel[2012]와 같은 선행연구에서도 소셜필터링은 협업필터링에 비해 계산 복잡도면에서 효율적이고, 예측정확도에 있어서도 협업필터링에 뒤지지 않으며, 신규 사용자에 대한 추천 정확도도 뛰어난 뿐 아니라 추천 사유를 효과적으로 설명할 수 있다는 점에서 장점이 많은 방법임을 주장한 바 있다. Avesani et al.[2004]의 연구에서는 소셜 정보가 협업필터링의 사용자-아이템 매트릭스의 희박성 문제를 보완함으로써 추천아이템의 커버리지를 증가시킨다고 하였다. Li and Kao[2009]는 소셜 네트워크의 신뢰에 기반한 추천시스템을 제안했는데 제안된 시스템이

기존 시스템에 비해 유의한 수준에서 정확도 향상을 가져온 것으로 보고하고 있다. Pitsilis and Knapskog[2009]은 제품에 대한 선호정보가 수집되지 않은 처음 사용자에게 관한 신규사용자 문제(Cold-start problem)를 해결하기 위해 기존 추천시스템에 신뢰정보를 결합한 하이브리드 추천기를 제안하였다. 무비렌즈(Movielens) 도메인을 대상으로 한 실험 결과, 제안한 하이브리드 추천기가 커버리지와 정확성 측면에서 협업필터링보다 근소한 차이로 우수함을 확인하였으며, 특히 신규사용자와 신규아이템에 대한 추천효과가 뛰어남을 발견한 바 있다.

이에 반해 소셜 정보가 추천시스템 성능 향상에 유의한 영향을 미치지 않는다는 연구도 다수 발견되고 있다. Golbeck et al.[2006]은 영화 정보를 추천하는 FilmTrust 사이트를 대상으로 사용자가 명시적으로 표현한 신뢰정보와 신뢰전파를 이용한 신뢰추론방법을 제안하였는데 실험결과 평균적인 제품 선호패턴과 동떨어진 사용자들에게는 신뢰기반 추론이 의미 있는 예측정확도를 나타내었지만 다수의 일반사용자에게 있어서는 신뢰기반 추천이 협업필터링 추천보다 더 효과적이지는 않은 것으로 보고하고 있다. Neto and Nowe[2012]는 협업필터링과 소셜 정보를 활용한 추천시스템을 비교하는 연구를 수행하였다. 이 연구는 예측정확도만이 아니라, 신규사용자에 대한 추천정확도, 계산 복잡성을 고려한 추천정확도, 설명능력 등 다양한 측면에서 추천시스템을 비교하였다는 점에서 의의를 가진다. 페이스북 어플로 개발된 마이팝콘(MyPopcorn)의 사용자를 대상으로 실험한 결과, 이웃의 수가 2 또는 3일 때는 협업필터링의 예측오차가 낮게 나타났으며 이웃의 수가 4 이상인 경우에는 소셜필터링의 예측오차가 낮게 나타난 것으로 보고하고 있다. Arazy et al.[2009], Ricci et al.[2011]과 Neto and Nowe[2012]는 소셜 정보에 기반한 추천시스템이

주목받는 연구토픽이 되고 있으며, 많은 연구자들이 소셜 추천시스템과 기존 추천시스템에 대한 성능비교 결과를 제시하고 있지만 아직 연구자들 사이에서 소셜 추천시스템의 성능이 더 우수하다는 사실에 대해서는 공감이가 형성되어있지 않은 상황이라고 언급하고 있다. 소셜 정보를 활용한 추천시스템에 대한 성능비교 연구 결과가 연구마다 다르게 나타나는 것은 어떤 도메인에서 어떤 정보를 어떤 방법으로 추천에 활용하는가의 차이 때문에 발생한다. 기존 소셜 추천시스템 연구의 경우, 단순 친구관계를 추천시스템에 활용하는 연구가 있는가 하면[Guy et al., 2009; Liu and Lee, 2010; Ma et al., 2008] 사용자가 명시적으로 표현한 신뢰정보를 추천에 활용하기도 하고[Ziegler and Lausen, 2004; Weng et al., 2006; O'Donovan, 2009; Zheleva et al., 2010; O'Doherty et al., 2012] 전파된 신뢰정보를 추천에 활용하기도 하며[Massa and Avesani, 2007a; Golbeck, 2009], 상호작용의 빈도나 관계 지속기간과 같은 관계강도를 추천에 활용하는[Arazy et al., 2009] 연구, 그리고 명성을 추천에 활용하는 연구[Avesani et al., 2004]도 있다. 또한 소셜 정보에 대한 개념정의도 연구자들마다 달라서 신뢰를 단순히 아는 관계(Knowing)로 해석하는[Golbeck, 2006] 연구에서부터 관계빈도나 친밀성 등으로 해석하는 연구, 사용자별 아이템에 대한 평가치의 유사성을 신뢰의 척도로 사용하는 연구[Guo et al., 2014] 등 다양한 형태로 소셜 정보가 정의되어 추천시스템에 활용됨으로써 연구자에 따라 추천시스템의 성능이 엇갈리게 보고되는 것으로 보인다. 따라서 본 연구에서는 다양한 소셜 정보를 몇 가지 형태로 정의하고, 동일한 조건 하에서 실험하여 소셜 정보의 추가적 활용이 추천시스템 성능향상에 효과적인지를 밝히고자 한다.

한편 두 번째 연구문제는 소셜 정보의 추가적인 활용이 추천 성능에 미치는 영향은 추천 도메

인에 따라 어떻게 달라질 것인지를 살펴보는 것이다. 기존 추천시스템 연구에서는 영화, 음악, 책과 같은 기호제품에서부터 전문정보나 서비스와 같은 높은 의사결정 위험이 수반된 제품에 이르기까지 연구자마다 서로 다른 도메인에서 얻어진 실험결과를 정리하여 소셜 정보 활용의 효과를 제시하였기 때문에 소셜 추천시스템의 성능에 대한 엇갈린 결과가 나타나기도 했다. 따라서 추천 도메인에 따라 소셜 정보의 활용이 추천시스템 품질에 얼마나 기여하는지를 살펴보는 것도 중요한 연구문제가 될 수 있다. Bonhard and Sasse[2006]는 추천 도메인을 계량화 및 측정이 가능한 제품을 대상으로 하는 객관적 도메인(Objective domain)과 개인의 기호제품을 대상으로 하는 기호 도메인(Taste domain)으로 구분하고, 객관적 도메인에 있는 제품으로 컴퓨터 하드웨어 및 소프트웨어, 전자제품, 자동차가 있고, 기호 도메인 제품으로 영화, 음악 등이 있음을 제시한 바 있다. 이 연구에서는 객관적 도메인의 경우 합리성에 기반하여 최종 의사결정이 이루어지는 경향이 있기 때문에 전문가 네트워크의 추천이 피추천인의 선택에 중요한 영향을 미친다고 하였다. 또한 기호 도메인에서는 추천인을 잘 알고 있는 경우, 추천인과의 기호유사도가 피추천인의 최종 판단에 중요한 영향을 미치며 추천인과의 과거경험이 없는 경우, 추천인에 대한 명성을 고려하여 피추천인의 최종 선택이 이루어진다고 하였다. 결국 Bonhard and Sasse[2006]의 연구는 단순 기호 제품의 추천에서는 선호유사도를 활용하는 협업필터링 추천이 효과적일 수 있으며, 신중한 선택이 요구되는 제품을 대상으로 하는 객관적 도메인일 경우 전문가와 같은 소셜 정보를 활용한 추천이 보다 효과적일 수 있음을 암시하고 있다. Huang et al.[2007]의 연구에서는 협업필터링 추천품질이 영화나 도서 판매사이트에서는 높지만 의류 판매사이트에서는

낮게 나타나는 것을 발견하고 그 이유를 비즈니스 형태나 거래데이터의 특성으로 기인한 것으로 보았다. 이상의 연구결과를 토대로 볼 때 의사결정의 위험이 얼마나 수반되는 도메인인지에 따라 소셜 정보가 추천품질에 미치는 영향은 다를 수 있음을 짐작할 수 있다. 기존 연구를 종합하면 소셜 정보에 대한 정의가 각 연구자마다 다르게 해석되어 사용되었을 뿐 아니라 추천시스템에 이용되는 소셜 정보도 다양하고, 실험대상 도메인도 상이하여 소셜 정보의 활용이 추천시스템 성능에 긍정적인 영향을 미치는지 비교가 어렵다. 본 연구에서는 추가적인 소셜 정보의 사용이 추천시스템의 성능을 개선할 수 있는지 그리고 어떤 도메인에서 소셜 정보를 추천에 활용하는 것이 추천 성능 향상에 효과적인지를 실증 분석을 통해 규명하고자 한다.

3. 용어정의

본 연구의 목표를 달성하기 위해서는 어떤 종류의 소셜 정보를 연구에 포함할지를 먼저 결정할 필요가 있다. 일반적으로 추천시스템에 활용되는 소셜 정보로는 사용자 간 신뢰관계, 명성, 팔로잉, 친구목록, 블로그 포스트, 댓글 및 의견, 태그(메타정보) 등이 있다. 이들 정보는 독립적으로 추천시스템을 구성하는데 활용될 뿐 아니라 서로 연계되어 추천시스템에 사용되기도 한다. 기존 연구를 분석하면, 소셜 추천시스템에서 주로 활용되어온 소셜 정보는 다음 <표 1>과 같이 요약된다. <표 1>에 제시된 소셜 정보 중 관계빈도는 친밀성에 대한 보다 구체화된 표현 중의 하나이고 태그 등의 메타 정보는 도메인에 따라 내용이 상이하다는 점, 그리고 유사성은 이미 기존 협업필터링에서 사용하고 있다는 점에서 이상의 세 가지 소셜 정보를 제외하고, 신뢰, 전문성, 친밀성, 관계지속성 정보를 선정하여 이들 정보의 추

〈표 1〉 소셜 추천시스템에 주로 활용되는 소셜 정보

소셜정보 유형	관련연구
명시적 또는 암묵적인 방법으로 유도된 사용자 간 신뢰관계 정보	Ziegler and Lausen[2004], Weng et al.[2006], Golbeck[2009], O'Donovan[2009], Zheleva et al.[2010], O'Doherty et al.[2012]
전문성이나 능력에 대한 타인의 평가를 객관적 지표로 나타낸 명성 또는 전문성 정보	Arazy et al.[2009], Song et al.[2013]
팔로잉, 지인, 친구관계 등의 친밀성 정보	Arazy et al.[2009], Neto and Nowe[2012], Li et al.[2013], Song et al.[2013], Bonhar and Sasse[2006]
사용자 속성이나 선호의 유사성	Groh and Ehmig[2007], Li et al.[2013], Song et al.[2013], Hu and Pu[2010], Bonhard and Sasse[2006]
포스트, 댓글 등의 상호작용 기록에서 유도된 관계강도 (관계지속성 및 빈도)	Arazy et al.[2009], Song et al.[2013]
태그 등의 메타정보	Tso-Sutter et al.[2008], Bogers and Bosh[2009], Gemmell et al.[2009], Gedikli and Jannach[2010]

가적인 활용이 추천시스템 성능향상에 얼마나 기여하는지를 살펴보고자 한다. 이중 전문성은 전문분야별로 존재할 수 있음을 고려하여 연구를 진행하기로 한다.

한편 Bonhard and Sasse[2006]의 연구와 Huang et al.[2007]의 연구에서 지적했듯이 의사결정의 위험이 얼마나 수반되는 도메인인지에 따라 소셜 정보가 추천품질에 미치는 영향은 다를 수 있음을 짐작할 수 있다. 관여이론(Involvement Theory)에 따르면 고관여(High Involvement) 제품의 경우, 소비자들은 제품정보의 탐색 및 구매 등의 단계에서 훨씬 많은 정신적 물리적 노력을 투자한다고 한다[Phang et al., 2010; Laaksonen, 1994]. 즉 고관여제품은 전자제품과 자동차와 같이 값이 비싸거나 자신에게 중요한 영향을 미치는 제품 또는 잘못 구매했을 때 많은 위험이 뒤따르는 제품을 의미하며, 구매할 때 여러 사람에게 물어보기도 하고 오랜 시간과 노력을 소비하면서 구매 과정에 깊이 관여하게 되는 특성을 가진다. 반면 저관여제품은 담배, 초콜릿과 같이 가격이 저렴하고 소비자가 구매를 잘못하더라도 위험성이 크지 않은 제품을 의미한다. 따라서 고관여제품을 구매할 때는 의사결정 과정, 정보처리 과정이

매우 복잡하게 되는 경향이 있다. 이러한 맥락에서 살펴보면 고관여제품일수록 소셜 정보에 의존하여 최종의사결정을 내리게 될 가능성이 많아진다. 그러므로 본 연구에서는 추천 도메인을 고관여제품 도메인과 저관여제품 도메인으로 구분하고 과연 도메인에 따라 소셜 정보가 추천 성능에 미치는 영향이 다르게 나타나는지를 살펴보기로 한다. 결국 본 연구에서는 다양한 소셜 정보를 신뢰, 전문성, 친밀성, 관계지속성 정보로 정의하고, 추천도메인을 제품관여 수준에 따라 고관여제품과 저관여제품 도메인으로 구분하여, 동일한 조건 하에서 실험하여 소셜 정보의 추가적 활용이 추천시스템 성능향상에 효과적인지를 밝히고자 한다.

4. 소셜 추천시스템 제안

4.1 소셜 추천시스템 설계

소셜 정보의 활용이 추천시스템 성능에 미치는 영향을 파악하기 위해서는 소셜 추천시스템을 구현할 필요가 있다. 일반적으로 추천시스템은 크게 협업필터링 추천, 콘텐츠기반 추천, 하이브리드 추천시스템으로 구분된다. 소셜 정보를

추천시스템에 활용하는 방식 또한 세 가지 추천 방식에 소셜 정보를 추가적으로 반영하여 추천을 할 수 있다. 그러나 콘텐츠 기반 추천과 하이브리드 추천 방식은 도메인에 따라 추천시스템 구성과 사용되는 정보가 달라지기 때문에 다양한 도메인을 대상으로 추천 성능을 측정 및 비교해야 하는 본 연구에는 적용이 곤란하다. 본 연구에서는 다양한 도메인에 공통적으로 적용될 수 있고 최근 활발한 상용화 적용이 이루어지고 있는 협업필터링 방식에 소셜 정보를 결합한 소셜 추천시스템을 제안하기로 한다. 여기서는 협업필터링 방식에 소셜 정보를 결합하여 추천을 할 수 있는 다양한 방식을 검토하고 본 연구에 필요한 효과적인 추천방법을 제시하기로 한다.

소셜 정보를 협업필터링방식의 추천과 결합하는 방법은 두 가지 방법이 있다. 첫 번째 방법은 소셜 정보를 이용한 추천과 협업필터링 방식에 의한 추천을 별개로 수행한 후 추천결과를 종합하는 방식이다[Claypool et al., 1999; Billsus and Pazzani, 2000]. 추천결과를 종합할 때는 예측된 추천점수를 가중 평균하는 방법으로 최종 추천 목록을 생성한다. 두 번째 방법은 협업필터링 방식에 소셜 정보를 동시에 고려하여 추천을 진행하는 방식이다. 즉 협업필터링 과정은 일반적으로 사용자별 아이템 선호 매트릭스로부터 사용자간 선호유사도를 계산하게 되는데, 유사도 계산과정에서 소셜관계 정보를 추가하여 유사도를 계산한 후 이웃을 구성하는 방식으로 추천을 진행할 수 있다. 이 방법은 Arazy et al.[2009], Massa and Avesani[2007b], Balabanovic and Shoham [1997], Hu and Pu[2010] 등 많은 연구자들에 의해 시도된 방법으로 본 연구에서도 이 방법을 이용하여 추천시스템을 구성하기로 한다. 두 번째 방식으로 소셜 추천시스템을 구성할 경우, 사용자간 유사도를 종합하는 방법이 필요하다. 이 방

법은 두 사용자 간 소셜관계에 의해 도출된 유사도 매트릭스와 제품 선호유사도 매트릭스를 종합하여 갱신된 유사도 매트릭스를 구하고 이 매트릭스에 의거하여 이웃을 선정한 후 최종 추천점수를 계산하는 방법이다. 이 때 유사도 매트릭스를 종합하여 갱신된 유사도 매트릭스를 구하는 방법은 Arazy et al.[2009]에서 제시한 방법을 적용하여 다음과 같은 방법으로 유도하기로 한다.

$$m_{u,v} = \begin{cases} s_{u,v}, & \text{소셜관계 유사도만 존재하는 경우} \\ p_{u,v}, & \text{선호 패턴 유사도만 존재하는 경우} \\ s_{u,v}, & \text{두 가지 유사도가 다 존재하는 경우} \end{cases} \quad (1)$$

$m_{u,v}$: 사용자 u 와 사용자 v 의 갱신된 유사도

$s_{u,v}$: 사용자 u 와 사용자 v 의 소셜관계 유사도

$p_{u,v}$: 사용자 u 와 사용자 v 의 선호 패턴 유사도

식 (1)에 나타난 선호패턴 유사도($p_{u,v}$)의 경우, 피어슨상관계수, 코사인계수, 로그우도 등의 다양한 유사도를 사용할 수 있으나 본 연구에서는 가장 일반적으로 적용되고 있는 피어슨상관계수를 이용하여 유사도를 계산하기로 한다. 최종 추천점수는 Resnick 공식[Resnick, 1994]을 적용하여 식 (2)와 같이 구할 수 있다. 식 (2)에서는 사용자마다 평가점수를 부여하는 성향이 다르기 때문에 사용자의 선호도 평가치 자체보다 평가치 편차($r_{v,i} - \bar{r}_v$)를 평가점수 계산에 활용하고 있다.

$$pr_{u,i} = \bar{r}_u + \frac{\sum_{v \in M} m_{u,v} (r_{v,i} - \bar{r}_v)}{\sum_v m_{u,v}} \quad (2)$$

$pr_{u,i}$: 사용자 u 의 아이템 i 에 대한 추천점수

$r_{v,i}$: 사용자 v 가 아이템 i 를 선호하는 정도

M : 갱신된 유사도에 의해 선정된 이웃사용자 집합

$m_{u,v}$: 사용자 u 와 사용자 v 의 갱신된 유사도

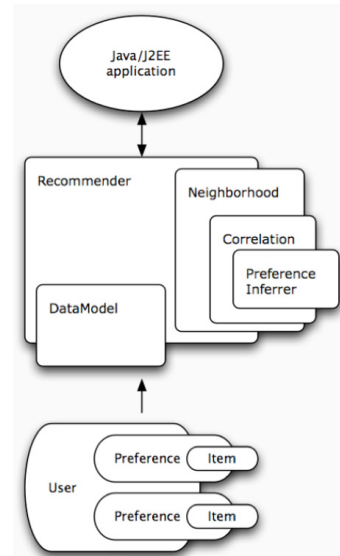
한편, 식 (2)에서 해당 사용자의 이웃이 없을 경우, 해당사용자가 다른 아이টে에 대해 평가한 값을 평균하여 ($pr_{u,i} = \bar{r}_u$) 추천점수를 계산하게 된다.

4.2 소셜 추천시스템 구현

본 연구에서는 소셜 추천기를 구현하기 위해 Apache Software Foundation에서 빅데이터용 오픈소스 프로젝트로 개발된 Mahout을 기반으로 구현한다. Mahout 추천기는 사용자기반 추천기와 아이টে기반 추천기로 구성되어 있으며 본 연구에서는 사용자 간 소셜 정보를 추천에 활용하기 때문에 사용자기반 추천기를 기반으로 추천시스템을 구성한다. 사용자기반 추천기는 <그림 1>에서와 같이 사용자별 아이টে 선호매트릭스를 보관하는 DataModel 컴포넌트와 사용자 간 선호패턴의 유사도를 계산하는 Correlation 컴포넌트, 그리고 도출된 유사도를 기반으로 이웃을 구성하는 Neighborhood 컴포넌트 등으로 구성되어 있다.

협업필터링 추천기의 경우 mahout에서 제공되는 기본 클래스들을 사용하여 구현이 가능하며, 소셜 추천기는 유사도(UserSimilarity) 인터페이스를 구현하는 클래스를 추가로 개발함으로써 구현이 가능하다. 아래 <그림 2>는 소셜

추천의 하나인 신뢰기반 추천기를 구현하기 위해 사용자기반 유사도 인터페이스를 구현한 SimilarityByTrustAndCF 클래스의 주요코드를 보여주고 있다.



<그림 1> Mahout의 사용자기반 추천기 프레임워크

5. 실험 결과

5.1 자료수집 방법

본 연구목적을 달성하기 위해서는 사용자별 제품 선호도에 대한 정보뿐 아니라 다양한 형태

<그림 2> SimilarityByTrustAndCF 클래스의 주요코드

```
public class SimilarityByTrustAndCF implements UserSimilarity {
    private HashMap<String, Integer> map = new HashMap<String, Integer>();
    private final DataModel dataModel
    @Override
    public double userSimilarity(long userID1, long userID2) throws TasteException {
        String idkey = Long.toString(userID1).concat(Long.toString(userID2));
        if (map.get(idkey) == null) {
            PearsonCorrelationSimilarity pcs = new PearsonCorrelationSimilarity(dataModel);
            return pcs.userSimilarity(userID1, userID2);
        }
        else return 1;
    }
}
```


의 사용자 간 소셜 관계에 대한 정보가 유지되는 온라인 공동체가 필요한데, 이 모든 정보가 관리되는 온라인 공동체는 현실에서 존재하지 않는다. 따라서 별도의 실험 집단을 구성하여 구성원별 제품 선호도 정보와 실험 집단 내 구성원 간 소셜 관계에 관한 자료를 수집하여 추천에 활용하였다. 실험 집단은 동일한 대학교에 재학 중인 학생 112명을 무작위 추출하여 구성하였다. 자료 수집은 구글독스를 기반으로 작성된 설문서를 이용하여 2014년 9월 15일부터 10일간 온라인 조사 방식으로 진행되었다. 112명의 학생 중 미 응답자와 불성실 응답자를 제외하고 최종 84명의 응답결과를 토대로 실험을 진행하였다. 응답학생은 대학 재학생으로 다양한 전공 및 학년으로 구성되어 있으며, 이중 38%는 여학생이었다.

실험을 위해 필요한 정보는 크게 사용자별 아이템 선호정보와 사용자 간 소셜 정보 두 가지로 구분된다. 먼저 사용자별 아이템 선호정보를 수집하기 위해 저관여제품으로 영화와 티셔츠를 선정하고, 고관여 제품으로 노트북 컴퓨터를 선정하고, 각 제품 도메인별로 25개의 아이템풀을 구성하고, 사용자별로 각 아이템에 대한 선호도를 표시하도록 하는 방식으로 조사를 진행하였다. 아이템에 대한 선호도는 대부분의 기존연구에서처럼 5점 척도(별점)로 조사하며 응답자가 잘 모르는 아이템에 대해서는 무응답을 할 수 있도록 하였다. 또한 소셜 정보는 실험집단 내 구성원간의 관계를 조사하는 것을 목표로 하여 각 응답자에게 실험집단 내에서 신뢰하는 친구, 친밀한 친구, 오래된 친구, 의류/티셔츠/노트북 분야별로 전문성이 있는 친구를 각각 2명씩 선택하도록 하여 조사를 하였다.

5.2 추천 성능 평가방법

일반적으로 추천시스템의 성능은 추천정확도, 추천품질, 커버리지, 참신성 등 다양한 측면에서 평가가 가능하다. 이중 추천정확도는 주로 (1-추

천오류율)로 계산하는데 추천오류율은 사용자가 평가한 선호도를 가린 후 추천시스템을 통해 선호도 예측치를 계산하고, 실제 사용자가 평가한 선호도와 예측 선호도의 차이를 이용하여 계산하는 것이 일반적이다. 이 때 실제 선호도와 예측 선호도의 단순차이를 이용하여 계산된 오류가 MAE (Mean Absolute Error)이며, 실제 선호도와 예측 선호도의 차이의 제곱으로 계산된 오류가 RMSE (Rooy Mean Square Error)이다. MAE와 RMSE가 평가집합(Test set)에 속한 전체 아이템을 대상으로 계산되는데 비해, 추천품질은 사용자에게 추천된 아이템 중 추천점수가 높은 Top n개의 아이템만을 대상으로 정확히 추천이 이루어졌는지를 판단하는 것이다. 추천품질을 평가하는 척도로는 정밀도(Precision)와 재현율(Recall) 그리고 F값이 있다. 정밀도(Precision)는 추천된 아이템들 (Top n) 중 실제 사용자가 선호하는 아이템의 비율을 의미하며, 재현율(Recall)은 사용자가 선호하는 아이템 중 Top n 추천리스트에 포함된 아이템의 비율로 계산된다. 또한 F값은 정밀도와 재현율을 종합하여 한 개의 척도로 구성한 것이다. 본 연구에서는 추천오류율(MAE, RMSE)과 추천품질(정밀도, 재현율, F)을 평가척도로 사용하여 소셜 정보의 추가적 활용이 추천시스템의 성능 향상에 미치는 영향을 파악하고자 한다.

추천품질 평가를 위해서는 먼저 특정사용자에 대해 아이템별로 예측된 추천점수가 적합한 추천인지 적합하지 못한 추천인지를 구분하는 기준이 필요하다. 본 연구에서는 Groh and Ehmig[2007]의 연구에서 사용한 방식을 수정 적용하여 식(3)과 같이 적합추천집합, 비적합 추천집합으로 구분하고 정밀도와 재현율을 각각 계산한다. 또한 사용자별로 예측점수가 높은 Top 5아이템을 추천하여 평가에 활용하기로 한다. 여기서 적합추천 아이템은 5점 만점에 예측 선호도가 3.0 이상인 아이템 중 실제선호도가 3.0 이상인 아이템이다.

$$\begin{aligned} pr_{u,i} \in P_a &\Leftrightarrow (r_{u,i} \geq 3 \wedge pr_{u,i} \geq 3) \\ pr_{u,i} \in P_{ia} &\Leftrightarrow (r_{u,i} < 3 \wedge pr_{u,i} \geq 3) \end{aligned} \quad (3)$$

P_a : 적합추천집합

P_{ia} : 비적합추천집합

각 척도의 계산식은 식 (4)와 같다. 전체 사용자별 선호값 중 70%를 추천기 생성에 사용하고 30%는 평가용 선호값으로 사용하였다.

$$MAE = \frac{1}{\|A\|} \sum_{(u,i) \in A} |pr_{u,i} - r_{u,i}|, \quad (4)$$

$$\text{where } A = \{(u,i) | pr_{u,i} \in P_a \cup P_{ia}\}$$

$$RMSE = \frac{1}{\|A\|} \sqrt{\sum_{(u,i) \in A} [pr_{u,i} - r_{u,i}]^2},$$

$$\text{where } A = \{(u,i) | pr_{u,i} \in P_a \cup P_{ia}\}$$

$$\text{Precision} = \frac{|P_a|}{|P_a| + |P_{ia}|}$$

$$\text{Recall} = \frac{|P_a|}{\|B\|}, \text{ where } B = \{(u,i) | r_{u,i} > 3\}$$

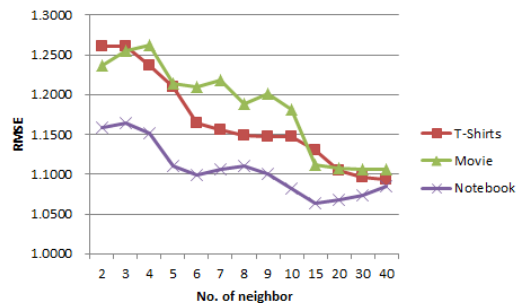
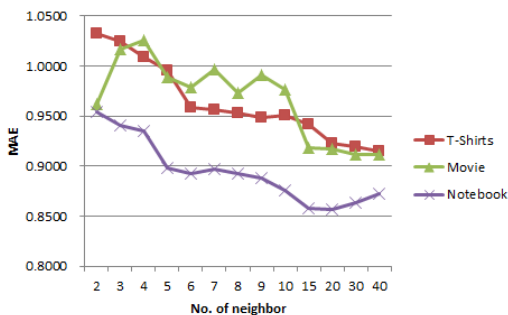
$$F = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

5.3 이웃수에 따른 추천 성능

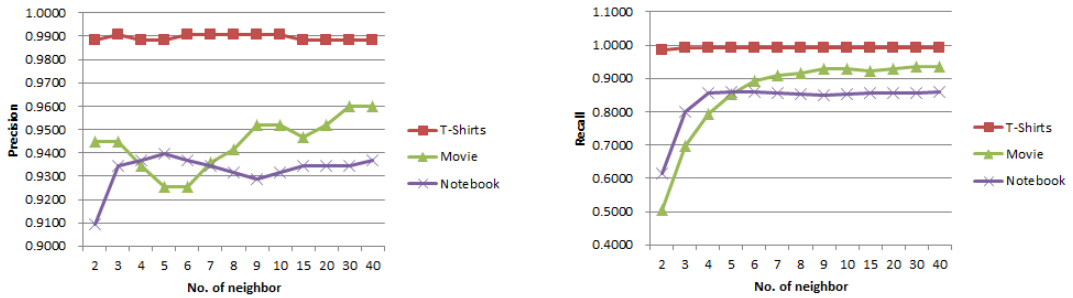
협업필터링 추천에서는 이웃의 수를 어떻게 결정하는가에 따라 추천 성능이 달라지는 경우

가 많다. 이웃의 수와 관련해서는 최적의 숫자라고 불리는 매직넘버가 존재하지 않는다는 것이 정설이다. 따라서 협업필터링 추천기의 성능을 평가할 때는 이웃의 수에 따라 추천 성능을 측정하여 가정 우수한 성능을 보이는 이웃의 수를 선택하여 실험을 진행하는 것이 일반적이다. 본 연구에서 제안한 소셜 추천기도 협업필터링 방식을 토대로 구성되었기 때문에 최적 이웃의 수를 결정하는 것이 필요하다. 여기서는 이웃의 수에 따라 추천오류율과 추천품질이 어떻게 달라지는지를 살펴보고 최적이웃을 결정하기로 한다. <그림 3>은 이웃의 수에 따른 제품별 추천오류율을 비교한 결과이다. 이웃의 수에 따른 추천오류율은 MAE와 RMSE 모두 이웃의 수가 증가하면서 대체로 낮아지는 경향을 나타내고 있다. 제품별 추천오류율(RMSE)은 티셔츠의 경우 이웃의 수가 40일 때, 영화의 경우 이웃의 수가 30일 때, 노트북의 경우 이웃의 수가 15일 때 가장 낮게 나타났다. 제품별 추천오류율은 티셔츠와 영화가 비슷하게 높은 오류율을 보이고 있으며, 노트북의 경우 티셔츠와 영화에 비해 현저히 낮게 나타났다.

<그림 4>는 이웃의 수에 따른 제품별 추천 품질을 비교한 결과이다. 정밀도(Precision) 측면에서 제품별 추천품질을 비교하면 티셔츠의 추천품질이 가장 높은 것으로 나타났고, 이웃의 수가



<그림 3> 이웃수에 따른 제품별 추천 오류율 비교

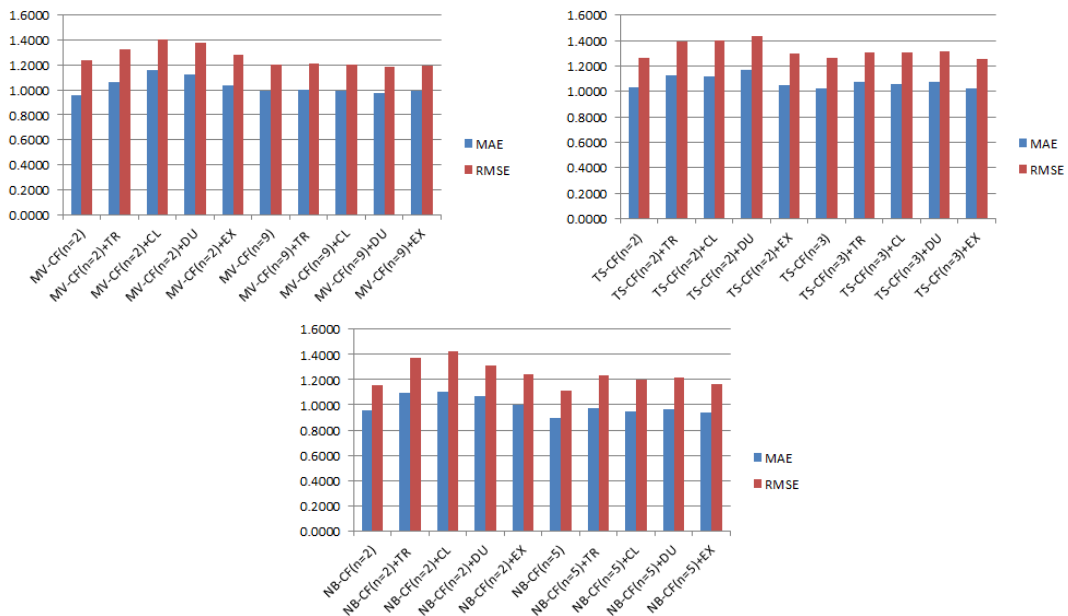


〈그림 4〉 이웃수에 따른 제품별 추천 품질 비교

4, 5, 6일 때를 제외하고는 영화의 추천품질이 그 다음으로 높았다. 재현율(Recall) 측면에서도 티셔츠의 추천품질이 가장 높았으며, 이웃의 수가 6 이상이 되면서 영화의 추천품질이 그 다음으로 높았다.

본 연구에서는 최적 이웃의 수를 선정하기 위해 제품별로 정밀도가 가장 높아지는 이웃의 수를 최적 이웃의 수로 선정하고 실험을 진행하기로 한다. 정밀도면에서는 티셔츠의 경우 이웃의 수가 3일 때, 영화의 경우 이웃의 수가 9 이상 일

때, 노트북의 경우 이웃의 수가 5일 때 가장 좋은 것으로 나타났다. 따라서 각 제품별, 이웃의 수를 티셔츠는 3, 영화는 9, 노트북은 5로 설정하여 추가적인 실험을 진행하기로 한다. 한편, 설문조사를 통해 소셜관계를 조사하는 과정에서 신뢰, 친밀성, 전문성, 관계지속성 등의 각 소셜관계에 대해 실험집단 내에서 2명을 선택하도록 조사했기 때문에 동일한 조건에서 비교를 위해 이웃의 수를 2명으로 설정한 실험결과도 동시에 제시하기로 한다.



〈그림 5〉 소셜 추천시스템과 협업필터링 추천시스템의 추천오류 비교

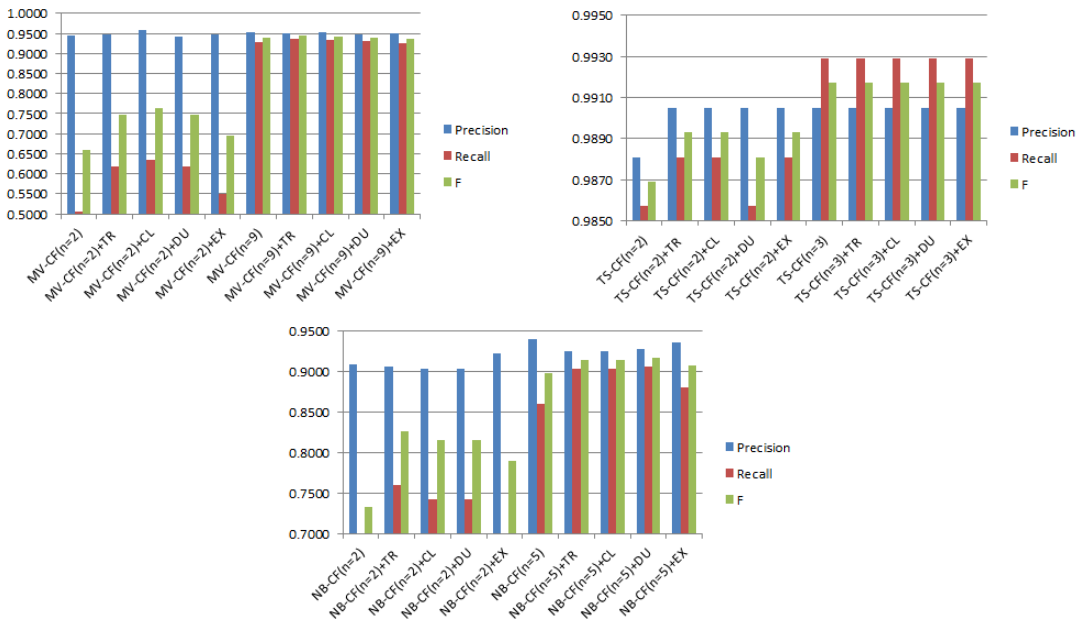
5.4 추천시스템에서 소셜 정보의 활용이 추천 성능에 미치는 영향

여기서는 소셜 정보의 활용이 추천 성능에 미치는 영향을 파악하기 위해 소셜 추천시스템과 기존 협업필터링 추천시스템을 추천오류율과 추천품질 측면에서 비교하기로 한다. <그림 5>의 상단 왼쪽 그래프는 영화도메인에서, 상단 오른쪽 그래프는 티셔츠 도메인에서, 하단 그래프는 노트북 도메인에서 각각 소셜 추천시스템과 협업필터링 추천시스템의 추천오류율을 비교한 결과이다.

영화의 경우, 이웃이 2일 때 기존 협업필터링 방식의 추천이 추천오류율이 가장 낮게 나타났으며(MAE = 0.9616, RMSE = 1.2363), 이웃이 9인 경우, 관계기간 정보를 추천에 활용했을 때 추천오류율이 가장 낮게 나타났다(MAE = 0.9742, RMSE = 1.1873). 티셔츠의 경우, 이웃이 2일 때 기존 협업필터링 방식의 추천이 추천오류율이 가장 낮게 나타났으며(MAE = 1.0330, RMSE =

1.2612), 이웃이 3인 경우, 전문성 정보를 추천에 활용했을 때 추천오류율이 가장 낮게 나타났다(MAE = 1.0209, RMSE = 1.2527). 한편 노트북의 경우, 이웃이 2일 때 기존 협업필터링 방식의 추천이 추천오류율이 가장 낮게 나타났으며(MAE = 0.9540, RMSE = 1.1581), 이웃이 5인 경우에도, 기존 협업필터링 방식의 추천오류율이 가장 낮게 나타났다(MAE = 0.8977, RMSE = 1.1103). 추천오류율 측면에서 소셜 정보의 활용효과를 살펴보면 영화와 티셔츠 같은 저관여제품 추천에서 관계지속성 또는 전문성 정보의 활용이 추천오류율을 낮추는 데 도움이 되었지만 대체로 기존 협업필터링방식에서 추천오류율이 오히려 낮게 나타남으로써 소셜 정보의 활용이 추천정확성에는 큰 영향을 미치지 못한 것으로 나타났다.

<그림 6>의 상단 왼쪽 그래프는 영화도메인에서, 상단 오른쪽 그래프는 티셔츠 도메인에서, 하단 그래프는 노트북 도메인에서 각각 소셜 추천시스템과 협업필터링 추천시스템의 추천품질



<그림 6> 소셜 추천시스템과 협업필터링 추천시스템의 추천품질 비교

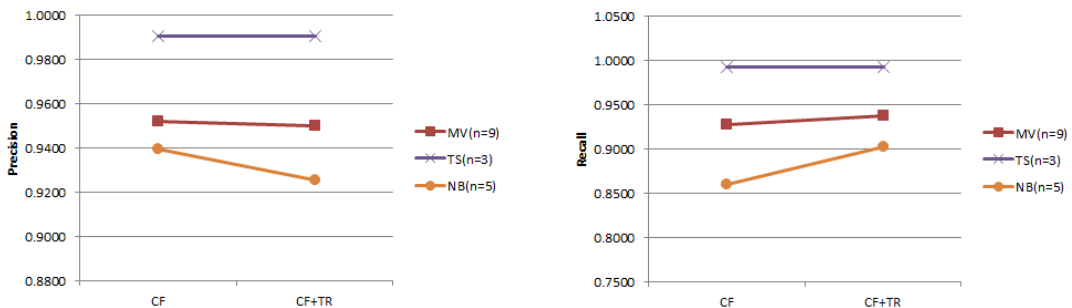
을 비교하고 있다.

영화의 경우, 이웃의 수에 따라 정밀도의 차이는 큰 변화가 없었으나 재현율은 민감하게 변화하였다. 이웃이 2인 경우, 정밀도와 재현율을 종합한 추천품질인 F값을 기준으로 비교할 때 기존 협업필터링 방식의 추천보다 소셜 정보를 활용한 추천시스템의 추천품질이 전반적으로 더 높게 나타났다. 특히 친밀성정보를 활용한 추천시스템의 추천품질($F = 0.7640$)은 기존 협업필터링 방식의 추천품질($F = 0.6594$)에 비해 15.8%나 더 향상된 것으로 나타났다. 이웃이 9인 경우에도 기존 협업필터링 방식의 추천품질($F = 0.9395$)에 비해 신뢰정보($F = 0.9438$), 친밀성정보($F = 0.9431$)를 활용한 추천시스템의 추천품질이 더 좋은 것으로 나타났다. 티셔츠의 경우에도 이웃의 수에 따라 재현율은 민감하게 변화하였다. 이웃이 2인 경우, 기존 협업필터링 방식의 추천보다 소셜 정보를 활용한 추천시스템의 추천품질이 전반적으로 더 높게 나타났다(신뢰 = 0.9893, 친밀성 = 0.9893, 전문성 = 0.9881, 기존CF = 0.9869). 이웃이 3인 경우, 기존 협업필터링 방식의 추천품질과 소셜 정보를 활용한 추천시스템의 추천품질이 공히 0.9917로 동일하게 나타났다. 노트북의 경우, 이웃이 2일 때는 기존 협업필터링 방식의 추천보다 소셜 정보를 활용한 추천시스템의 추천품질이 전반적으로 더 높게 나타났다(신뢰 = 0.8270, 친밀성 = 0.8156, 관계기간 =

0.8155, 전문성 = 0.7897, 기존CF = 0.7327). 특히 신뢰정보를 활용한 추천시스템의 품질은 기존 협업필터링 방식의 추천시스템에 비해 추천품질이 12.8% 향상되었다. 이웃이 5인 경우에도 소셜 정보를 활용한 추천이 기존 협업필터링 방식의 추천보다 추천품질이 더 높았다(신뢰 = 0.9143, 친밀성 = 0.9143, 관계기간 = 0.9169, 전문성 = 0.9076, 기존CF = 0.8985). 결과적으로 추천품질이 유사했던 티셔츠(이웃 = 3)추천의 경우를 제외하고는 모든 경우에서 소셜 정보를 활용한 추천이 기존의 협업필터링 추천방식보다 높은 추천품질을 보여주었다. 소셜 정보를 활용한 추천시스템의 추천품질이 높은 것은 정밀도의 차이는 미미한데 비해 재현율의 차이가 현저했기 때문이었다.

5.5 관여수준에 따른 소셜관계가 추천 성능에 미치는 영향

여기서는 소셜 정보의 활용이 추천품질에 미치는 영향이 제품관여 수준에 따라 어떻게 달라지는지를 살펴보기로 한다. 먼저 <그림 7>은 신뢰정보의 활용이 추천품질에 미치는 영향이 제품 도메인별로 어떻게 달라지는지를 보여주고 있다. <그림 7>에서 왼쪽 그래프는 신뢰정보의 활용이 정밀도에 미치는 영향을 제품별로 본 것이며, 오른쪽 그래프는 신뢰정보의 활용이 재현율에 미치는 영향을 제품별로 본 것이다.

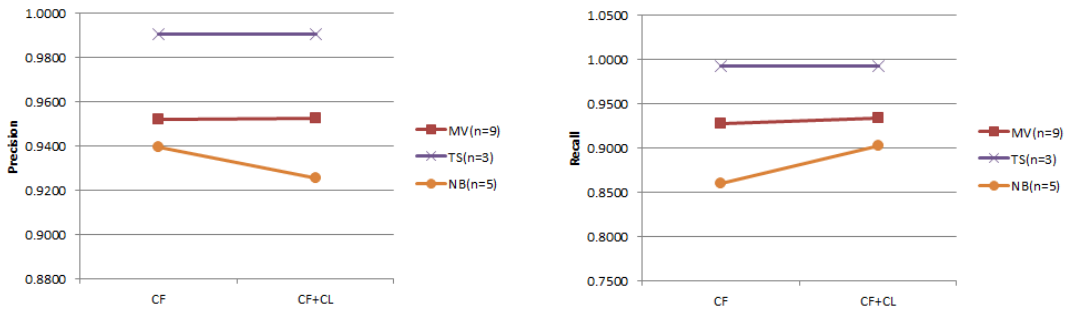


<그림 7> 제품도메인별 신뢰정보의 활용에 따른 추천품질의 변화

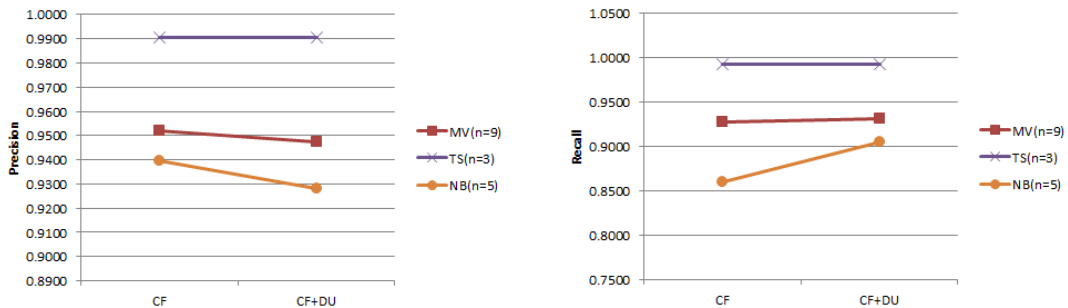
제품도메인별 신뢰정보의 활용에 따른 추천품질의 변화를 살펴보면, 영화나 티셔츠 같은 저관여제품에서 보다는 노트북과 같은 고관여제품에서 신뢰정보의 활용이 추천품질에 더 많은 영향을 미치는 것으로 나타났다. 노트북의 경우, 신뢰정보를 추천에 활용함으로써 기존 협업필터링 추천에 비해 정밀도에 있어서는 낮아졌으나 재현율은 많이 높아졌다. 그러나 영화와 티셔츠의 경우에는 신뢰정보를 활용한다고 해서 추천품질의 변화가

거의 없었다(그래프의 기울기가 0에 가까움). 즉 영화나 티셔츠의 경우, 신뢰정보를 활용하여 추천을 하더라도 기존 협업필터링 방식으로 추천했을 때와 비교하여 추천품질의 변화가 미약하였으나 노트북의 경우, 신뢰정보를 활용할 경우 추천품질이 더 민감하게 달라지는 것으로 나타났다.

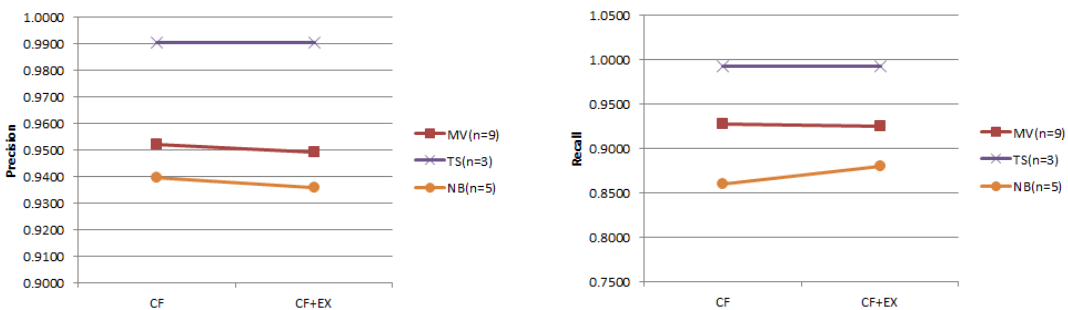
<그림 8>~<그림 10>은 각각 제품도메인별로 친밀성정보, 관계지속성정보, 전문성정보의 활용이 추천품질에 미치는 영향을 보여주고 있다. 세



<그림 8> 제품도메인별 친밀성정보의 활용에 따른 추천품질의 변화



<그림 9> 제품도메인별 관계지속성정보의 활용에 따른 추천품질의 변화



<그림 10> 제품도메인별 전문성정보의 활용에 따른 추천품질의 변화

그림에서 왼쪽 그래프는 소셜 정보의 활용이 정밀도에 미치는 영향을 제품별로 본 것이며, 오른쪽 그래프는 재현율에 미치는 영향을 제품별로 본 것이다. 제품 도메인별로 친밀성, 관계지속성, 전문성정보의 활용에 따른 추천품질의 변화는 신뢰정보의 활용 때와 유사하게 나타났다. 즉 영화나 티셔츠 같은 저관여제품에서 보다는 노트북과 같은 고관여제품에서 소셜 정보의 활용이 추천품질에 더 많은 영향을 미치는 것으로 나타났다.

6. 결론

본 연구에서는 추가적인 소셜 정보의 사용이 추천품질을 개선할 수 있는지 그리고 어떤 도메인에서 소셜 정보를 추천에 활용하는 것이 추천 성능 향상에 효과적인지를 실증 분석을 통해 규명하였다. 추천오류율 측면의 실증분석 결과, 영화와 티셔츠 같은 저관여제품 도메인에서 관계지속성 및 전문성 정보의 활용이 추천오류율을 낮추는 데 도움이 되었지만 대체로 기존 협업필터링방식에서 추천오류율이 오히려 낮게 나타남으로써 소셜 정보의 활용이 추천정확성에는 큰 영향을 미치지 못한 것으로 나타났다. 반면에 추천품질 측면에서는 티셔츠(이웃 = 3)추천의 경우를 제외하고는 모든 경우에서 소셜 정보를 활용한 추천이 기존의 협업필터링 추천방식보다 높은 추천품질을 보여주었다. 소셜 정보를 활용한 추천시스템의 추천품질이 높은 것은 정밀도의 차이는 미미한데 비해 재현율의 차이가 현저했기 때문이었다. 한편, 제품도메인별로 소셜 정보의 활용에 따른 추천품질의 변화를 살펴보면, 영화나 티셔츠 같은 저관여제품에서 보다는 노트북과 같은 고관여제품에서 공통적으로 소셜 정보의 활용이 추천품질에 더 많은 영향을 미치는 것으로 나타났다. 즉 영화나 티셔츠의 경우, 소셜

정보를 활용하여 추천을 하더라도 기존 협업필터링 방식으로 추천했을 때와 비교하여 추천품질의 변화가 미약하였으나 노트북의 경우, 소셜 정보를 활용할 경우 추천품질이 더 민감하게 달라지는 것으로 나타났다.

본 연구의 한계는 다음과 같다. 첫째는 추천정확도 평가 방법과 관련된 문제를 지적할 수 있다. 본 연구는 소셜 정보를 사용자 간 유사도 계산에 활용함으로써 추천을 하는 사전적인 방법을 사용하였다. 즉 나와 관계를 맺고 있는 구성원이 선호하는 아이템을 나에게 추천해 주는 방식으로 추천이 이루어졌다. 그러나 나와 관계를 맺고 있는 구성원이 추천한 것을 내가 좋아할 수는 있지만 그 구성원이 선호하는 것과 내가 선호하는 것이 항상 유사한 것은 아니다. 따라서 보다 바람직한 추천 성능 평가를 위해서는 실험집단 내 구성원들에게 추천자를 알려주면서 직접 추천을 수행하고 추천 만족도를 조사하는 방식으로 이루어진다면 소셜 정보가 추천 성능 향상에 미치는 영향은 본 실험결과보다 더 긍정적인 결과로 나타날 것이다. 두 번째 한계는 자료수집의 어려움으로 인해 다양한 제품도메인과 다양한 사용자 계층을 대상으로 자료를 수집하지 못하였다는 점을 들 수 있다. 본 연구에서는 고관여제품으로 노트북컴퓨터를 저관여제품으로 영화와 티셔츠를 선정하여 실험을 하였는데 세 가지 제품 도메인만으로 결과를 일반화하기에는 부족한 점이 존재한다. 또한 사용자 측면에서도 대학생만을 대상으로 자료를 수집하였기 때문에 다양한 사용자 계층으로 확대해석하기에는 한계가 있다. 지금까지 연구에서는 다양한 소셜 정보의 추가적인 활용이 추천 성능 향상에 어떻게 영향을 미치는지에 대한 종합적인 비교연구는 찾아볼 수 없다. 본 연구에서는 상기에 제시된 여러 한계점에도 불구하고 기존에 상반된 주장이 제기된 소셜 정보가 추천시스템 성능향상에 얼마나 기여

할 수 있는지를 동일한 조건에서 실증연구를 통해 파악하였다는 점에서 의의를 가진다. 본 연구의 한계를 극복하여 추천시스템에서 소셜 정보의 가치를 규명할 수 있는 보다 일반화된 후속연구가 향후 과제로 필요하다. 본 연구결과는 제품 도메인에 따라 어떤 소셜 정보를 결합하여 추천 시스템을 구성하는 것이 효과적인지를 제시하였기 때문에 다양한 개인화 추천시스템을 설계할 때 설계 가이드라인으로 활용될 수 있을 것이다.

참 고 문 헌

- [1] Arazy, O., Kumar, N., and Shapira, B., "Improving Social Recommender Systems", *Journal of IT Professional*, Vol. 11, No. 4, 2009, pp. 31-37.
- [2] Avesani, P., Massa, P., and Tiella, R., "Moleskiing : A Trust-Aware Decentralized Recommender System", *Proceedings of the First Workshop on Friend of a Friend Social Networking and the Semantic Web*, 2004.
- [3] Balabanovic, M. and Shoham, Y., "Content-based, collaborative recommendation", *Communications of the ACM*, Vol. 40, No. 3, 1997, pp. 66-72.
- [4] Billsus, D. and Pazzani, M., "User modeling for adaptive news access", *User Modeling and User-Adapted Interaction*, Vol. 10, No. 2, 2000, pp. 147-180.
- [5] Bogers, T. and Bosch, A., "Collaborative and content-based filtering for item recommendation on social bookmarking websites", *Proceedings of the 2009 ACM Conference on Recommender Systems*, 2009, pp. 9-16.
- [6] Bonhard, P., "Who do trust? Combining recommender systems and social networking for better advice", *International Conference on Intelligent User Interfaces*, 2005.
- [7] Bonhard, P. and Sasse, M. A. "Knowing me, knowing you : Using profiles and social networking to improve recommender Systems", *BT Technology Journal*, Vol. 24, No. 3, 2006, pp. 84-98.
- [8] Carrer-Neto, W., Hernandez-Alcaraz, M. L., Valencia-Garcia, R., and Garcia-Sanchez, F., "Social knowledge-based recommender system : Application to the movies domain", *Expert Systems with Applications*, Vol. 39, No. 12, 2012, pp. 10990-11000.
- [9] Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., and Sartin, M., "Combining content-based and collaborative filters in an online newspaper", *Proceedings of ACM SIGIR Workshop on Recommender Systems*, 1999, pp. 40-48.
- [10] Gedikli, F. and Jannach, D., "Rating items by rating tags", *Proceedings of the 2010 ACM Conference on Recommender Systems*, 2010, pp. 25-32.
- [11] Gemmell, J., Schimoler, T., Ramezani, M., Christiansen, L., and Mobasher, B., "Improving FolkRank with item-based collaborative filtering", *Proceedings of the 2009 ACM conference on Recommender Systems*, 2009, pp. 17-24.
- [12] Gilly, M., Graham J., Wolfenbarger M., and Yale L., "A Dyadic Study of Interpersonal Information Search", *Journal of the Academy of Marketing Science*, Vol. 26, 1998, pp. 83-100.
- [13] Golbeck, J., "Generating Predictive Movie Recommendations from Trust in Social

- Networks”, *Proceedings of the Fourth International Conference on Trust Management*, 2006.
- [14] Golbeck, J., “Trust and nuanced profile similarity in online social networks”, *ACM Trans.*, Web, Vol. 3, No. 4, 2009.
- [15] Groh, G. and Ehmig, C., *Recommendations in taste related domains : collaborative filtering vs. social filtering*, (Gross and Inkpen, K. ed.), ACM, 2007.
- [16] Guo, G., Zhang, J., Thalmann, D., Basu, A., and Yorke-Smith, N., “From Ratings to Trust : an Empirical Study of Implicit Trust in Recommender Systems”, *Symposium on Applied Computing- Recommender Systems*, 2014.
- [17] Guy, I., Zwerdling, N., Carmel, D., Ronen, I., Uziel, E., Yogev, S., and Ofek-Koifman, S., “Personalized recommendation of social software items based on social relations”, *Proc. of RecSys '09*, 2009, pp. 53-60.
- [18] Hu, H. R. and Pu, P., “Using personality information in collaborative filtering for new users”, *Proceedings of the 2010 ACM Conference on Recommender Systems*, 2010, pp. 17-24.
- [19] Huang, Z., Daniel D. Z., and Hsinchun C. “Analyzing consumer-product graphs : Empirical findings and applications in recommender systems”, *Management science*, Vol. 53, No. 7, 2007, pp. 1146-1164.
- [20] Kim, H. N., Alkhalidi, A., Saddik, A. E., and Jo, G. S., “Collaborative user modeling with user-generated tags for social recommender Systems”, *Expert Systems with Applications*, Vol. 38, No. 7, 2011, pp. pp. 8488-8496..
- [21] Laaksonen, P., *Consumer involvement : concepts and research*, Routledge, London, 1994.
- [22] Levin, D. and Cross R., “The Strength of Weak Ties You Can Trust”, *Management Science*, Vol. 40, No. 11, 2004, pp. 1477-1490.
- [23] Li, Y. M. and Kao, C. H. P., “TREPPS : a trust-based recommender system for peer production services”, *Expert Systems with Applications*, Vol. 36, No. 2, 2009, pp. 3263-3277.
- [24] Liu, F. and Lee. H. J., “Use of social network information to enhance collaborative filtering performance”, *Expert Systems of Applications*, Vol. 37, No. 7, 2010, pp. 4772-4778.
- [25] Ma, H., Yang, H., Lyu, M. R., and King, I., “SoRec : Social recommendation using probabilistic matrix factorization”, *Proc. of CIKM '08*, 2008, pp. 931-940.
- [26] Massa, P. and Avesani, P., “Trust metrics on controversial users : balancing between tyranny of the majority and echo chambers”, *International Journal on Semantic Web and Information Systems*, Vol. 3, No. 1, 2007a, pp. 39-64.
- [27] Massa, P. and Avesani, P., “Trust-aware recommender systems”, *Proc. of RecSys '07*, 2007b, pp. 17-24.
- [28] Neto, W. L. M. and Nowe, A., “Insights on Social Recommender System”, *ACM Recommender Systems*, 2012.
- [29] O'Doherty, D., Jouili, S., and Van Roy, P., “Trust-Based Recommendation : An Empirical Analysis”, *Sixth ACM Workshop on Social Network Mining and Analysis (SNA-*

- KDD 2012), 2012.
- [30] O'Donovan, J., "Capturing trust in social web applications", *Computing with Social Trust*, Vol. 3, 2009, pp. 213-257.
- [31] Phang, C. W., Kankanhalli, A., Ramakrishnan, K. and Raman, K. S., "Customers' preference of online store visit strategies : an investigation of demographic variables", *Proceedings of EJIS*, 2010, pp. 344-358.
- [32] Pitsilis, G. and Knapkog S. J., "Social trust as a solution to address sparsity-inherent problems of recommender systems", *ACM RecSys*, 2009.
- [33] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., and Riedl, J., "GroupLens : An Open Architecture for Collaborative Filtering of Netnews", *Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*, 1994, pp. 175-186.
- [34] Ricci, F., Rokach, L., and Shapira, B., "Introduction to recommender systems handbook", *Recommender Systems Handbook*, 2011, pp. 1-35.
- [35] Sarwar, B., Karypis, G., Konstan, J., and Riedl, J., "Analysis of Recommendation Algorithm for E-Commerce", *ACM SIG EC*, 2000.
- [36] Smith, D., Menon, S. and Sivakumar K., "Online peer and editorial recommendations, trust, and choice in virtual markets", *Journal of Interactive Marketing*, Vol. 19, No. 3, 2005, pp. 15-37.
- [37] Song, H. S., "Prediction Method for the Implicit Interpersonal Trust Between Facebook Users", *Journal of Information Technology Applications and Management*, Vol. 20, No. 2, 2013, pp.177-191.
- [38] Tso-Sutter, K. H. L., Marinho, L. B., and Schmidt-Thieme, L., "Tag-aware recommender systems by fusion of collaborative filtering algorithms", *Proceedings of the 2008 ACM Symposium on Applied Computing*, 2008, pp. 1995-1999.
- [39] Weng, J., Miao, C., and Goh, A., *Improving collaborative filtering with trust-based metrics*, (In H. Haddad and H. Haddad, ed.), SAC ACM, 2006, pp. 1860-1864.
- [40] Zheleva, E., Getoor, L., Golbeck, J., and Kuter, U., "Using friendship ties and family circles for link prediction", *Proceedings of the Second international conference on Advances in social network mining and analysis, SNAKDD '08*, 2010, pp. 97-113.
- [41] Zheng, N. and Li, Q., "A recommender system based on tag and time information for social tagging systems", *Expert Systems with Applications*, Vol. 38, No. 4, 2011, pp. 4575-4587.
- [42] Ziegler, C.-N. and Lausen, G., *Analyzing correlation between trust and user similarity in online communities*, (In Jensen et al., ed.), 2004, pp. 251-265.

■ 저자소개



송희석

고려대학교 경영학과에서 학사, 한국과학기술원에서 석사 및 박사학위를 취득하였다. 대우전자와 대우정보시스템에서 15년간 근무하였고 현재 한남대학교 경영정보학과 교수로 재직 중이다. Artificial Intelligence Review, Knowledge-based Systems, Expert Systems, Expert Systems With Applications 등 다양한 국내외 저널에 논문을 출판하였다. 관심분야는 CRM과 Data Mining, 유비쿼터스 비즈니스, 비즈니스모델, 소셜 네트워크 등이다.



주석정

아주대학교 화학공학과에서 학사, 연세대학교에서 석사, 가톨릭대학교에서 박사학위를 취득하였다. 시스템공학연구소, 대성산업((주), 삼일회계법인, 중소기업기술정보진흥원에서 근무 하였고 현재 한남대학교 경영정보학과 교수로 재직 중이다. Information Systems Review, Journal of Information Technology and Decision Making, 중소기업연구, 정보시스템연구, 산업정보학회지 등 국내외 전문학술지에 논문을 출판하였다. 주요 관심분야는 ERP, AIS, ISP, 프로젝트 관리, 중소기업정보화 정책 등이다.



이재훈

건국대학교 법학과에서 학사, 경북대학교에서 경영학 석사, 금오공과대학교에서 기술경영 컨설팅학 박사학위를 취득하였다. SK Broadband와 CJ Hellovision에서 12년간 근무하였고, 현재 창업진흥원 스마트창업팀에서 공공데이터 분야 지원과 ICT 스마트 동향보고서를 매주 발행하며 근무 중이다. 청년창업의도에 개인적 특성, 대학 및 사회적 지원적 특성이 미치는 영향에 관한 연구, 대학생 창업활동에 관한 실증 연구 등 기술창업 및 청년창업분야에 다수의 논문을 발표하였다. 관심분야는 창업기업 전략 및 기업가정신 등이다.