

협업필터링시스템의 발전 및 활용 방안 연구*

김진목†

강남대학교 산업데이터사이언스학부

본 연구의 목적은 개인 또는 유사한 정보요구를 갖고 있는 그룹에 대한 맞춤형 정보서비스를 제공할 수 있는 협업필터링 시스템의 발전을 고찰하고, 이를 바탕으로 그 활용 방안을 모색하기 위함이다. 본 연구의 목적을 달성하기 위해서 첫째, 협업필터링을 포함한 정보필터링의 현황과 발전을 고찰하고 둘째, 다양한 형태의 협업필터링 시스템 사례를 조사하고 셋째, 협업필터링 시스템의 문제점을 파악하고 마지막으로, 협업필터링 시스템의 활용 방안을 제시하였다. 맞춤형 정보서비스는 인공지능 및 사물인터넷(IoT)의 발전과 더불어 더욱 중요하게 인식되고 있으며 정보이용, 상품구매, 서비스제공 등 다양한 분야에서 활용되고 있다. 본 연구의 결과는 보다 효율적인 맞춤형 정보서비스의 활용 방안을 제시하는데 그 의의가 있다고 하겠다.

주요어: 정보필터링, 협업필터링, 맞춤형 정보서비스, 정보검색

* 본 연구는 2016학년도 강남대학교 교내연구비 지원에 의하여 수행된 연구임.

† 단독저자 : 김진목, 강남대학교 산업데이터사이언스학부 부교수, 경기도 용인시 기흥구 강남로 40 천은관 510호,
E-mail : kjinmook@kangnam.ac.kr

■ 최초투고일 : 2018년 3월 1일 ■ 심사마감일 : 2018년 3월 27일 ■ 게재확정일 : 2018년 3월 27일

1. 서론

정보필터링 시스템은 이용자가 필요로 하는 정보를 제공한다는 점에서 정보검색 시스템과 유사하다. 일반적으로, 정보검색은 정보를 필요로 하는 이용자가 검색을 통하여 시스템으로부터 정보를 추출한다는 의미에서 “풀(pull)”서비스라 지칭할 수 있다. 반면, 정보필터링은 시스템이 이용자의 요구를 미리 파악해서 이용자 맞춤형 정보를 추천한다는 의미에서 “푸쉬(push)”서비스라 지칭할 수 있다(Kim, Oard & Romanik, 2000). 즉, 정보필터링 시스템은 이용자의 정보요구를 파악해서 맞춤형 정보를 제공해 준다는 점에서 정보검색 시스템과 차이가 있다고 하겠다.

정보필터링 시스템은 문헌의 콘텐츠로부터 추출된 정보 즉, 단어 또는 구문 등을 기반으로 정보를 검색한다는 점에서 정보검색 시스템의 일종이라고 할 수 있다. 위에서 언급한 정보검색 시스템과는 차별화된 이용자 맞춤형 정보를 제공하기 위해서는 그들의 정보 요구를 미리 파악하는 것이 필수적이라 하겠다. 이러한 이유로 대다수 정보필터링 시스템은 사용자들의 정보요구를 파악하여 관리하는 사용자 프로파일을 생성·유지·관리하고 있다. 정보검색 시스템에 새로운 문헌 또는 아이템이 유입되면 이용자 프로 파일에 저장된 개별 이용자의 정보요구와 비교하여 적합한 정보라고 판단될 경우 추천하게 된다.

이용자가 원하는 정보를 시스템이 선제적으로 탐색하여 제공할 수 있다는 점에서 정보필터링 시스템은 최근 온라인 마케팅을 포함한 여러 분야에서 활용되고 있다. 특히, 인공지능 및 사물인터넷의 발전과 더불어 수많은 정보 속에서 개별 이용자가 필요로 하는 정보를 찾아 주는 맞춤형 정보 서비스는 그 중요성이 더욱 커지고 있다. 예를 들면, 이용자의 정보요구를 파악하여 필요한 정보를

제공함으로써 이용자 만족도를 높인다거나, 고객의 수요를 예측하여 제품 또는 아이템을 추천해주는 서비스 등을 들 수 있다.

오늘날 많은 이용자 또는 소비자들은 검색엔진, 소셜 네트워크 서비스(SNS), 온라인 쇼핑을 포함한 다양한 형태의 온라인 활동을 통해서 그들의 정보요구 또는 수요를 유추할 수 있는 대량의 데이터를 생산하고 있다. 인공지능을 활용한 이용자 프로파일 구축 및 빅데이터 분석을 통해서 보다 정교한 이용자 맞춤형 정보서비스를 제공할 수 있는 기술적 기반과 인프라가 구축되어 있으며, 이를 활용한 정보 또는 고객서비스 분야는 더욱 발전하리라 예상된다.

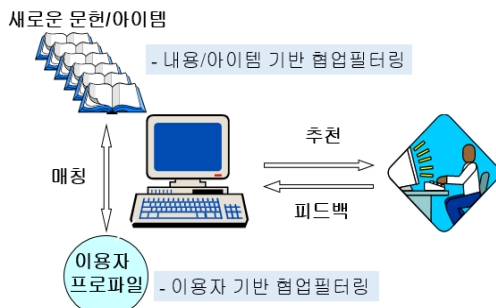
본 연구는 개인별 또는 정보요구를 공유하는 그룹별 맞춤형 정보서비스를 제공할 수 있는 정보필터링 시스템 특히, 정보요구를 공유하는 이용자들이 상호 협력하여 정보를 제공받을 수 있는 협업필터링 시스템의 발전 및 활용 방안에 대하여 고찰해 보기 위함이다. 이를 위해, 2장에서는 협업필터링 시스템의 현황을 조사하고, 실제 응용 사례를 조사하여 협업필터링 시스템이 어떻게 활용되고 있는지 살펴보고, 선행문헌 분석을 통해서 협업필터링의 발전을 살펴보았다. 3장에서는 협업필터링 시스템 구축 시 고려하여야 할 잠재적 문제점을 살펴보고 그 해결 방안을 논의하였다. 결론에서는, 앞에서 논의한 내용을 바탕으로 보다 효율적인 맞춤형 정보서비스 제공을 위한 협업필터링 시스템의 활용 방안을 제시하였다.

2. 현황, 사례 조사 및 선행연구 분석

1) 현황

정보필터링은 1990년대 인터넷의 등장과 더불어

어 유즈넷(USENET) 뉴스와 같은 그 당시 대량의 정보로부터 관심 주제의 정보를 선별하기 위한 목적으로 활용되기 시작하였다(Stevens, 1993). 초기의 정보필터링 시스템은 문헌에 포함된 단어 등에 기반을 두어 유사 문헌을 탐색하는 콘텐츠 기반 필터링 시스템(content-based filtering system)이 주를 이루었으나(Sheth, 1994), 이후 특정 주제 분야에 관심이 있는 이용자들의 평가에 기반을 두어 적합문헌을 추천해 주는 협업필터링 시스템(collaborative filtering system or recommender system)이 등장하였다(CACM, 1997).



<그림 1> 협업필터링 시스템의 종류

협업필터링 시스템은 <그림 1>에 제시된 바와 같이 내용 또는 아이템 기반 협업필터링과 이용자 기반 협업필터링으로 구분할 수 있다. 내용 또는 아이템 기반 협업필터링은 위에 기술한 콘텐츠 기반 필터링 시스템과 마찬가지로 문헌 또는 아이템으로부터 파생된 단어나 특징을 이용하여 유사 문헌 또는 아이템을 탐색한다. 반면에, 사용자 기반 협업필터링은 문헌 또는 아이템의 내용이나 특징을 이용하지 않고 특정 문헌을 이미 읽은 이용자나 특정 아이템을 사용한 경험이 있는 사용자들의 선호도 또는 만족도 피드백을 통해서 추천이 이루어진다. 즉, 특정 문헌 또는 아이템에 대한 선호도 또는 만족도 및 이용자 간의 유사도

를 활용하여 새로운 문헌 또는 아이템을 추천하는 방식을 이용자 기반 협업필터링이라고 할 수 있다.

1990년대 웹의 등장과 더불어 텍스트 형식의 정보자료와 더불어 이미지, 동영상 등 멀티미디어 자료가 폭발적으로 증가해 왔다. 또한, 온라인에서 이루어지는 아이템, 상품, 서비스 구매 등의 행위는 정보의 내용을 기반으로 검색하는 전통적인 방식으로는 이용자의 다양한 요구를 충족시키기 어렵게 되었다. 이용자 기반의 협업필터링은 콘텐츠가 아닌 이용자의 사용 또는 구매 경험에 따른 선호도를 정보 추천의 메커니즘으로 사용하기 때문에 주제적합성을 주로 이용하는 기존의 방식으로는 검색이 어려운 새로운 형식의 정보를 검색하기에 보다 적합한 방식으로 주목 받기 시작했다.

2) 사례 조사

협업필터링에 기반을 둔 추천시스템은 그룹렌즈(Konstan et al., 1997), 무비렌즈(Miller et al., 2003) 등 영화를 대상으로 한 초기의 연구 이후 다양한 분야에서 활발한 연구가 이루어져왔다. 협업필터링 시스템의 활용 및 발전 방안에 대한 시사점을 고찰하고자, 본 절에서는 실생활에서 많이 이용되는 영화, 음악, 도서, 쇼핑 분야의 추천시스템 사례를 조사하여 정리하였다.

(1) 영화

국외의 경우, 2000년을 전후해서 영화 추천시스템에 관한 연구가 활발하게 이루어 졌으며 2000년대 후반기부터 상용화되어 이용되기 시작하였다. 1997년에 설립된 넷플릭스(Netflix)는 동영상 스트리밍 서비스를 제공하며 시네매치(Cinematch)

라는 영화추천 시스템을 개발하여 사용해 왔다. 이용자가 선택한 영화, 별점(<그림 2> 참조) 등의 데이터를 수집·분석하여 이용자 개개인의 취향에 맞는 영화를 추천해준다.

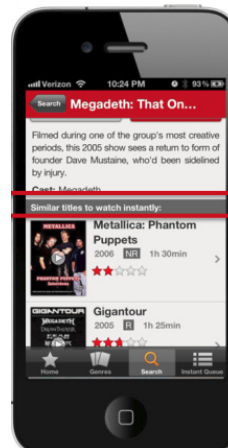
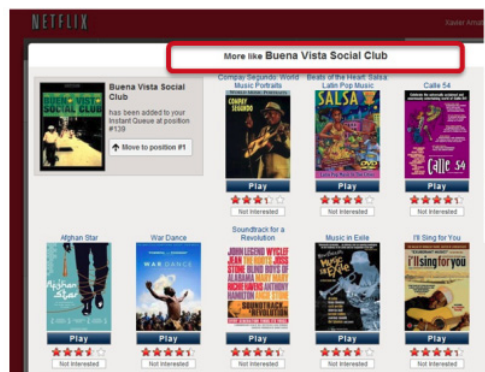
2006년 넷플릭스는 임의로 선정된 100만개의 영화 평점(rating) 데이터를 연구자들에게 제공하고 기존의 추천시스템 성능 향상을 위해 연구자들을 대상으로 공개 콘테스트를 진행하였다. 공개 콘테스트에 참여한 연구자들은 기계학습 및 데이터마이닝 기법을 적용하여 시네매치의 성능을 10% 이상 향상시킬 수 있었다(Bennett, Lanning, 2007). 이를 통해, 넷플릭스는 보다 정확하게 개인의 취향을 예측할 수 있었으며, 이는 이용자 만족도 향상이라는 결과로 나타나게 되었다. 넷플릭스 이외에도 지니(Jinni), 테이스트 키드(Taste Kid), 나노 크라우드(Nanocrowd), 로튼 토마토(Rotten Tomatoes), IMDb, 클리티커(Criticker) 등 다수의 웹사이트에서 협업필터링을 적용한 영화 추천서비스를 제공하고 있다.

국내의 경우, 프로그램스(Programs)라는 벤처 기업이 개발한 왓차(Watcha)라는 영화 추천서비

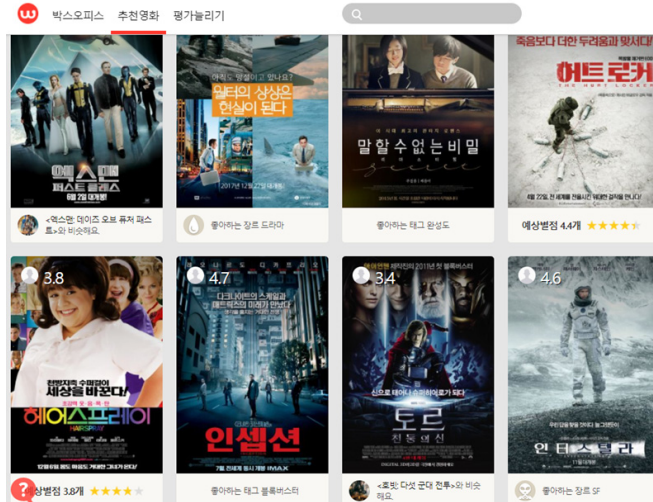
스를 들 수 있겠다(<그림 3> 참조). 영화에 대한 이용자의 선호도 또는 취향을 분석한 뒤 새로운 영화에 대한 이용자 선호도를 예측하여 별점을 제공하고 추천해주는 서비스이다(이종설 외, 2017). 왓차는 기계학습 및 데이터 분석을 통한 개인화 기술을 기반으로 하고 있으며 2012년 서비스 제공을 시작한 후 비교적 단기간에 네이버보다 많은 영화 별점을 수집하였다. 현재는, 영화분야 이외에도 드라마, 다큐멘터리, 애니메이션, 도서 등 다양한 분야로의 확장을 도모하고 있다(박태훈, 이충재, 2014).

(2) 음악

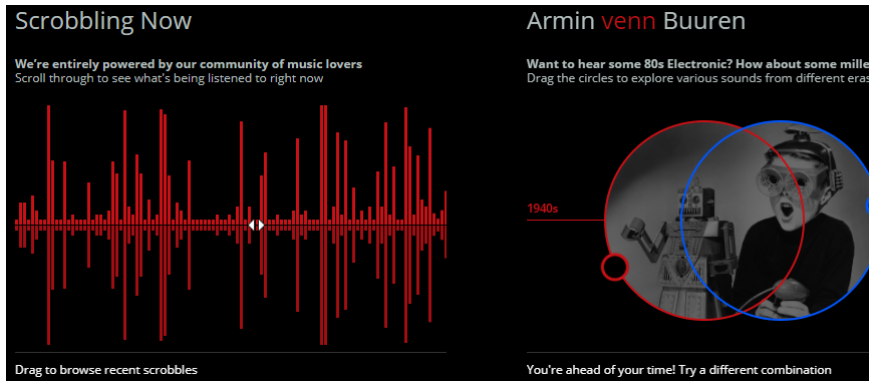
인터넷의 발달과 더불어 영화와 마찬가지로 음악은 온라인상에서 많은 이용자들이 각자의 취향에 맞는 음악을 구매하여 들을 수 있게 되었다. 개별 이용자들의 취향에 맞는 음악을 추천하여 그들의 음악에 대한 구매 요구를 충족시키기 위해서 협업필터링에 기반을 둔 음악 추천시스템에 관한 연구가 이루어져왔다.



<그림 2> 넷플릭스 영화 추천
(출처: 넷플릭스 홈페이지)



<그림 3> 왓차 추천 영화
(출처: 왓차 홈페이지)



<그림 4> 라스트 FM 오디오 스크로블러
(출처: 라스트에프엠 홈페이지)

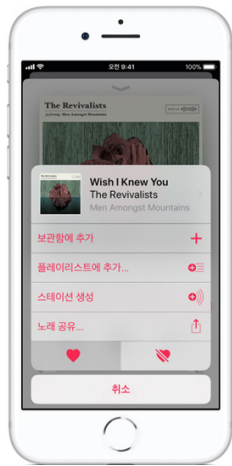
영국의 인터넷 라디오 방송 라스트 FM(Last.fm)은 <그림 4>에 제시된 바와 같이 음악 추천시스템인 오디오 스크로블러(Audioscrobber)를 통해 이용자의 음악 청취 행태를 분석하여 취향에 맞는 음악을 추천한다. 또한, 개별 이용자의 음악 취향을 파악하여 취향이 유사한 이용자들을 선호도를 바탕으로 서로 연결시켜 주고 이용자 개개인 맞춤형의 라디오 스테이션을 제공해 준다

(Haupt, 2010).

판도라(Pandora)는 이용자가 'like'라고 평가한 음악의 특색을 바탕으로 해당 사용자의 프로파일을 구축 한 뒤, 신곡이 출시되면 그 곡을 분석하여 장르, 비트, 음색 등 약 400여 항목의 특성을 추출한다. 이용자 프로파일과 분석한 곡의 특성을 비교함으로써 이용자의 취향에 부합하는 음악을 추천한다. 또한, 이용자 프로파일을 토대로 계

절, 상황 등에 맞는 새로운 음악을 추천하기도 한다(Abdollahpouri & Essinger, 2017).

라스트 FM, 판도라 이외에도 음악을 추천하는 또 다른 서비스로 협업필터링 및 딥러닝 기술을 사용하는 스포티파이(Spotify), 애플 뮤직(Apple Music), 멜론 등의 서비스를 들 수 있다. 국내에서도 서비스를 제공하는 애플 뮤직은 이용자가 노래 또는 앨범을 청취할 때 그 곡에 대한 선호 여부를 <그림 5>의 예시와 같이 알려 주고 애플 뮤직은 이용자 피드백을 바탕으로 더 적합한 음악을 추천해 준다.



<그림 5> 애플 뮤직 ‘좋아요’ 또는 ‘별로예요’ 선택하기
(출처: 애플 뮤직 홈페이지)

(3) 도서

도서는 웹에서 많은 이용자들이 구입하는 상품 중 하나이며 아마존(Amazon)을 비롯한 대다수 온라인 서점들이 협업필터링의 개념을 적용하여 소비자 마케팅을 적용해 왔다. 구매자가 구입한 특정 도서를 구입한 다수의 다른 이용자들이 구매한 책을 추천하는 것이 그 예라하겠다. 현재, 아마

존은 도서 이외에도 전자제품, 의류, 뷰티, 아웃도어 등 다양한 상품 및 서비스를 포함하는 온라인 쇼핑몰로 발전하였으며, 상품의 추천을 위해 몇 가지 방법으로 사용자의 성향을 파악한다. 제품에 대한 별점, 구매 이력 등 대부분 사용자가 해당 페이지에 접속한 후 활동한 내역들을 관찰하여 파악한 데이터에 기반하고 있다.

국내에서는 전자책 서비스업체인 리디북스가 <그림 6>에 제시된 바와 같이 빅데이터 개념을 활용한 ‘맞춤추천’ 서비스를 통해 도서를 추천하고 있다. 2012년 도입된 ‘나를 위한 추천책’으로부터 시작된 ‘맞춤추천’은 독자의 도서 구매 이력, 도서 검색 및 클릭 정보 등의 정보를 바탕으로 독자가 좋아하는 분야를 분석해 도서를 찾아 추천해 주는 개인 맞춤형 서비스이다. 리디북스가 자체적으로 개발한 이 서비스는 카테고리, 작가 등의 메타 정보들을 이용한 ‘콘텐츠 기반추천 방식’과 많은 사용자들로부터 얻은 기호정보에 따라 사용자의 관심사들을 자동적으로 예측하게 해주는 ‘협업적 필터링 방식’이 모두 적용된다. 이를 통해 독자들의 선호도와 관심 표현을 나타내는 데이터를 분석해 아직 구매하지 않은 도서들을 독자들에게 추천해주고 있다.



<그림 6> 리디북스 맞춤추천 서비스
(출처: 리디북스 홈페이지)

(4) 쇼핑

트윈슈머는 쌍둥이를 지칭하는 트윈(Twin)과 소비자(Consumer)의 합성어로 생각, 취미, 기호, 반응, 소비 등의 성향이 마치 쌍둥이들처럼 유사한 사람들과 온라인으로 연결되어 그들의 의견을 듣고 행동에 옮기는 소비자들을 일컫는 용어이다. 즉, 인터넷에서 제품 사용 후기, 소비자 또는 전문가 추천 등을 참고하여 구입 여부를 판단하는 소비자 층을 의미한다. 일례로, 많은 온라인 소비자들이 가전제품 등의 구입 시 다른 소비자의 사용 후기를 참고하며 그들 또한 직접 사용 후기를 작성하기도 한다(김지영 외, 2008).

트윈슈머들을 대상으로 한 마케팅에서는 제품에 대한 취향, 기호 등 서로 간의 상관성 파악이 중요하며, 이를 위해 협업 필터링 기술이 이용될 수 있다. 세계적인 정보기술 시장조사업체인 포레스터리서치에 따르면 유럽 소비자가 가전제품을 구매할 때 50퍼센트 이상이 제품 사용 후기를 참고해서 구매하며, 30퍼센트 이상은 제품 사용 후기에 매겨진 평가점수에 근거해서 구매하고, 15퍼센트 이상은 자신이 직접 사용 후기를 작성하는 것으로 나타났다(Forrester Research, 2008).

국내의 경우에도 특정 제품군을 평가하고 후기를 올리는 전문사이트가 있다, 디씨인사이드(디지털카메라), 세티즌(스마트폰), 모니터포유(모니터) 등은 한국형 트윈슈머의 예들이다. 사용 후기를 적용한 추천시스템은 위에 기술한 영화, 음악, 도서는 물론 맛 집, 상품, 여행 서비스 등 대부분의 온라인 쇼핑에 적용되어 사용되고 있다.

3) 협업필터링

위에서 논의한 바와 같이 협업필터링 시스템은 인터넷의 등장, 온라인 쇼핑의 괄목할 만한 성장,

인공지능의 발달 등으로 인해서 2000대 이후 빠르게 상용화 되었으며 향후 맞춤형 정보서비스 제공을 위한 활용도가 더욱 증대할 것이 예상된다. 보다 효율적인 협업필터링 시스템의 활용방안을 논의하기 위해 1절에서는 협업필터링 시스템의 현황과 발전을 고찰하였고, 2절에서는 실제 활용 사례를 영화, 음악, 도서, 쇼핑 분야를 중심으로 살펴 보았다. 협업필터링 시스템에 관한 보다 심층적인 고찰을 위해 본 절에서는 협업필터링 시스템에 관한 국내와 국외의 연구를 조사·분석하였다.

쿠마르, 벤바셋(2006)은 추천과 소비자 리뷰가 온라인 쇼핑 사이트 소비자들에게 주요한 역할을 한다는 현상에 주목하고, 이를 실증적으로 입증하기 위해 유용성과 사회적 실재감의 변수를 사용한 연구를 수행하였다. 아마존닷컴의 콘텐츠를 실시간으로 필터링하여 단순히 고객제품 리뷰가 있고 없고의 차이가 사용자들의 사회적 실재감에 미치는 영향을 조사하였으며, 소비자 리뷰의 존재만으로 사회적 실재감이 유의하게 증가하는 것으로 확인하였다. 추천과 소비자 리뷰를 제공하는 것이 그 웹사이트의 유용성과 사회적 실재감을 향상시킨다는 연구결과를 제시하였다.

김경재, 안현철(2009)은 추천시스템에 가장 많이 활용되고 있는 추천기법 중 하나인 협동필터링이 가질 수 있는 희박성과 확장성의 문제를 완화하면서도 추천의 성과를 유지 또는 높일 수 있는 보완된 협동필터링을 제안하고자 연구를 수행하였다. 유전자 알고리즘을 활용한 사용자-상품 매트릭스의 두 차원, 즉 사용자 차원과 상품 차원의 효과적인 축소기법을 제시하여 희박성과 확장성 문제를 개선해 보고자 시도하였다. 또한, 협동필터링의 효율성 개선과 더불어 적절한 사용자-상품만 유사도 계산에 사용하여 예측정확도를 높이고자 하였다. 실험결과, 제안 모형들이 전통적인 협동필터링 기법과 비교하여 보다 효율적이고

더 높은 정확도를 산출함을 확인하였다.

무담비, 셔프(2010)는 상품 또는 서비스에 대한 온라인 소비자 후기(평가)가 빠르게 증가하고 있으며 그 중요성이 더욱 증대하고 있다는데 주목하고, 소비자 후기(평가)의 유용성 평가 모델을 개발하였다. 아마존(Amazon.com)에서 판매 하는 6개 제품에 대한 1,587개의 소비자 후기(평가)를 분석한 결과, 평가의 극단성, 깊이, 제품 종류 등이 소비자 평가의 유용성에 영향을 미친다고 주장하였다. 제품의 종류에 따라서 평가의 극단성이 소비자 후기(평가)의 유용성에 미치는 영향을 완화시켜 주었으며, 구매 경험이 있는 제품의 경우에는 극단적인 평가보다 중도적인 평가가 더 유용하다는 것을 밝혔다. 제품의 종류, 평가의 깊이는 소비자 평가의 유용성에 긍정적인 영향을 미쳤으나, 제품의 종류는 평가의 깊이가 유용성에 미치는 영향을 감소시키기도 하였다고 주장하였다.

이석환, 박승현(2011)은 협력적 필터링 시스템에서 추천의 정확도를 향상시키기 위한 두 가지 방법: 1) 기존의 선호도 예측방법을 개선한 이웃 사용자 중심의 선호도 예측방법; 2) 연령 속성을 반영한 추천순위 결정방법을 제시하였다. 이웃 사용자 중심의 선호도 예측방법은 목표고객의 이웃 사용자가 입력한 선호도에 기반하여 선호도를 예측함으로써 그 정확도를 높일 수 있었다. 또한, 예측 선호도와 더불어 목표고객과 이웃사용자의 연령 속성을 추천순위 결정에 반영함으로써 보다 정확한 추천목록을 생성할 수 있었다. 결과적으로, 연구에서 제시한 두 가지 방법 모두 기존 방식을 사용한 추천 실험결과에 비해 추천의 정확도를 높일 수 있었음을 확인하였다.

안성만, 김인환, 최병구, 조윤희, 김은홍, 김명균(2012)은 협업필터링 추천과 관련하여 적용 도메인에 따라 추천 성과 간에 차이가 다르게 나타난다는 점에 주목하고, 소셜네트워크의 구조적 측정

지표를 활용하여 추천 성과 간의 차이가 발생하는 원인을 체계적으로 규명하고자 하였다. 이를 위해 소셜네트워크의 구조적 측정지표와 협업필터링 추천 성과 간의 관계에 대한 가설을 수립하고 국내 H백화점의 거래데이터를 활용하여 이를 실증적으로 검증하고자 연구를 수행하였다. 검증 결과 밀도와 포괄성은 추천 성과에 긍정적인 영향을 미치는 반면 군집화계수는 부정적인 영향을 미치는 것을 파악하였다.

한소월, 이민수(2012)는 소셜 네트워크에 존재하는 대량의 데이터를 분석·활용한 추천기법에 적합한 데이터 모델을 제안하고자 연구를 수행하였다. 사람들이 온라인 커뮤니티, 블로그, 트위터 등 다양한 소셜미디어를 통해 자신의 견해나 의견을 표현하고 정보교환을 한다는 점에 주목하고, 이를 추천 시스템에 적용함으로써 개개인의 취향과 성향에 부합하는 추천을 할 수 있다고 주장하였다. 또한 사용자들에게 제공하는 추천 정보의 신뢰도 및 품질개선에도 기여를 할 수 있다고 하였다. 대량의 소셜 데이터를 저장하고 처리하기 위해서는 그에 맞는 데이터 모델과 처리 방식이 필요하며, 이를 위해 빅데이터 기반의 데이터 모델을 제안하였다. 또한, 검증을 위하여 사용자의 특성과 취향에 맞는 신뢰도가 높은 추천기법을 적용하여 데이터 모델의 적용가능성을 실험하였다.

이석필(2013)은 최근의 미디어 생성/소비 패턴은 UCC 같이 소비자가 직접 미디어를 생성하고 소비하는 프로세스가 등장하여 일반화되고 있다는 점에 주목하고, 이를 이용해 맞춤형 정보서비스를 할 수 있는 추천 에이전트를 개발하였다. 즉, 다종의 멀티미디어 콘텐츠 소비를 바탕으로 사용자들을 소셜 네트워크화하고, 콘텐츠 선호도가 유사한 사용자들의 프로파일을 바탕으로 맞춤형 콘텐츠를 추천하였다.

정중희, 김종우(2013)는 기존 협업필터링 기법에

서 주로 사용하는 고객 간의 선호도 유사성과 사회 네트워크 척도의 접목을 통한 추천 성능 향상 방안을 제시하였다. 유사성을 기초로 구성된 고객 간의 네트워크 척도인 중심성을 활용하여 특정 고객을 위한 적합한 상품 정보를 선정토록 하고, 이를 기존 협업필터링의 추천 성과와 비교하였다. 유튜브(youtube.com) 방문 로그데이터를 활용한 실증적 비교 결과, 유사성과 중심성의 접목을 통해서 추천 성과를 향상시킬 수 있음을 입증하였다.

이오준, 유은순(2015)은 모델 기반 협업필터링은 희박성 문제와 확장성 문제 등 협업 필터링의 한계를 개선하지만, 모델 생성 비용이 높고 성능/확장성 트레이드오프가 발생한다는 또 다른 한계점을 갖는다는데 주목하였다. 이를 해결하기 위해, 군집화 기반 협업 필터링에 마르코프 전이확률모델과 퍼지 군집화의 개념을 접목한 예측적 군집화 기반 협업 필터링 기법을 제안하였다. 실험결과, 제안된 기법이 의미 있는 수준의 성능 향상을 나타내지는 않았으나 성능 안정성의 측면에서는 기존의 기법들을 개선하고 있다고 주장하였다.

최슬비, 광기영, 안현철(2016)은 협업필터링을 통한 대다수 추천 시스템들이 익명의 사용자가 평가한 점수에 근거하여 추천결과를 생성하는 한계점이 있음에 주목하고, 자신이 신뢰하는 타인의 추천을 활용할 수 있는 방안을 제시하였다. 이를 위해, 소셜 네트워크분석에서 네트워크 내의 중심적 위치를 보여주는 척도인 내향 및 외향 중심성을 활용하여 사용자 간 유사도를 산출하는 알고리즘과 사용자 신뢰 네트워크를 탐색하여 추천 대상이 되는 사용자가 직접·간접적으로 신뢰하는 사용자의 평가점수를 보다 높게 반영하는 알고리즘을 제안한 뒤 그 성능을 비교하였다. 연구 결과, 사용자 신뢰 네트워크를 기반으로 탐색하는 알고리즘이 가장 우수한 성능을 보였으며, 기존 협업 필터

링에 비해서 통계적으로 유의한 수준의 정확도 개선이 이루어졌음을 확인하였다.

유석중(2017)은 최근 온라인 상거래에서 개인의 성향을 고려한 맞춤형 상품 추천을 위해 사용되고 있는 협업필터링 기법이 확장성, 희박성, 다양성 등 문제점을 갖고 있음에 주목하였다. 이러한 문제점들의 해결 및 협업필터링을 통한 추천 정확도 개선을 위해 신뢰도 기반 유사도 평가 모델을 제안하였다. 이웃들과의 평점 오차만을 고려하는 기존 유사도 평가 모델과 달리, 제시된 모델은 아이템 인기도를 신뢰도로 변환하여 예측 선호도의 정확도를 개선하고자 하였다.

선행연구 분석결과, 협업필터링은 인공지능의 발전과 더불어 추천의 정확도를 개선하고자 하는 연구와 더불어 추천의 확장성, 희박성, 다양성 등 협업필터링이 갖고 있는 문제점들을 해결하기 위한 연구가 수행되었다. 온라인 쇼핑이 일반화되고 더 많은 소비자들이 웹에서 정보를 교환하고 상품 및 서비스를 구매하는 추세이기에, 협업필터링을 통한 추천시스템들은 향후 더 많은 분야에서 활발하게 활용되리라 예상된다. 인공지능을 적용한 추천의 정확도 향상 및 협업필터링의 문제점을 완화하기 위한 이 분야의 연구 또한 지속적으로 수행되어야 할 것이다.

3. 협업필터링의 문제점

이용자 기반 협업필터링을 통한 추천은 충분한 수의 이용자들로부터 선호도 파악을 위한 충분한 양의 데이터를 수집해야 비로소 추천이 가능해진다. 즉, 이용자들 간의 선호도 파악을 위한 초기 데이터 수집의 어려움이 있다. 초기 데이터 수집의 어려움은 이용자 기반 협업필터링에 콘텐츠 기반 필터링을 접목하여 어느 정도 해결이 가능

할 것이다.

또 다른 문제점으로는, 이용자 프로파일의 지속적인 갱신에 많은 노력이 필요하다는 것이다. 추천의 정확도를 높이기 위해 이용자들의 요구, 취향 선호도 등을 파악하여 이용자 프로파일을 구축하고 이를 유지·관리·갱신하는 노력이 지속적으로 이루어져야 한다. 초기에는 이용자들로부터 직접 피드백을 수집하여 이용자 프로파일을 구축하고, 이를 기반으로 구매, 이용, 후기작성 등 이용자들의 정보행태를 암묵적으로 관찰하여 이용자 프로파일을 갱신할 수 있을 것이다.

이외에도, 익명의 이용자로부터 제공되는 추천에 대한 신뢰와 책임의 문제, 사용자와 상품의 증가로 인한 데이터 양의 방대함에서 기인하는 확장성의 문제, 추천 결과의 편향성 및 회소성 등 협업필터링이 갖고 있는 근본적인 문제점들이 있다. 이러한 문제점들을 해결하려는 노력을 통해서 협업필터링 시스템의 활용이 더욱 활성화 될 수 있을 것이다.

4. 결론: 활용 방안

웹에서 정보의 교환은 물론 상품과 서비스를 구매하는 사람들이 빠르게 증가하고 있다. 정보환경의 변화는 정보 이용자들이 더 이상 이용자의 역할에만 머물지 않고 UCC의 예와 같이 정보의 생산자이면서 동시에 간단한 태깅, 이용 후기, 댓글 등을 쓰면서 정보의 처리자/조직자역할을 하기도 한다. 영화, 음악, 도서, 전자제품 등 상품의 구매 및 서비스와 관현해서는 사용자 리뷰가 상품의 구매 여부에 지대한 영향을 미치게 되었다.

정보환경의 변화와 더불어서 빅데이터로 대변되는 데이터·정보의 양적 증가는 기존의 정보검색 시스템이 처리할 수 있는 실질적 범위를 능가하게

되었다. 양적인 증가는 물론, 트위터, 페이스북 등 소셜 네트워크 서비스를 통해서 쏟아져 나오는 비정형 데이터들은 기존의 관계형 데이터베이스 관련 기술로는 적시에 수집·처리·분석하기가 사실상 불가능 하게 되었다. 반면에, 더욱 더 많은 수의 사용자들이 소셜 네트워크 서비스를 통해서 자신들의 취향, 의견 등을 적극적으로 표현하기 시작하였고, 소비자의 요구와 취향을 반영하여 마케팅에 활용하고자 하는 비즈니스적 요구와 맞물려서 협업필터링을 통한 추천시스템들이 각광을 받게 되었다.

USENET 뉴스의 선별을 위해 시작된 정보필터링은, 오늘날 사용자들의 유사도 및 선호도를 정보·상품·서비스의 추천에 활용하는 협업필터링의 개념으로 발전하게 되었다. 협업 필터링은 정보의 주제를 뛰어 넘어 유사한 이용자의 선호도를 기반으로 정보를 추천하므로, 주제만을 기반으로 정보를 찾아주는 전통적인 정보 검색시스템이 다루기 힘든 멀티미디어·상품·서비스 등 새로운 형태의 정보자료를 효과적으로 찾아 줄 수 있는 대안으로 부상하였다.

본 연구에서는 협업 필터링 시스템의 보다 효과적인 활용 방안을 모색해 보고자, 협업필터링의 현황과 발전을 고찰하고, 영화, 음악, 도서, 트윈슈머 등 협업 필터링의 실제 활용 사례를 살펴보았다. 또한, 협업 필터링 분야의 선행연구들을 조사·분석하여 협업필터링을 통한 추천의 정확도 향상, 이용자 프로파일 구축의 어려움, 확장성 및 회소성 등 협업 필터링의 문제점들을 논의하였다.

4차 산업혁명과 더불어 각광을 받고 있는 인공지능과 빅데이터 분석 기술, 사물인터넷을 통해 실시간으로 쏟아져 나오는 센서 데이터, 소셜 네트워크 서비스를 통한 이용자들이 생산하는 비정형 데이터의 폭발적 증가 등은 협업 필터링 시스템의 유용성 및 활용도를 더욱 증대시키고 있는

현실이다. 알파고로 대변되는 인공지능 기술의 괄목할 만한 성장은 이용자들 간의 유사도 및 선호도를 더욱 정교하게 파악하게 함으로써 협업필터링을 활용한 추천의 정확도를 높일 수 있게 되었다. 그러므로, 협업필터링을 통한 추천 시스템은 정보의 교환, 온라인 쇼핑, 소비자 마케팅 등 더욱 더 많은 분야에서 활용될 것임은 쉽게 예상해 볼 수 있겠다.

그러나, 위에서 논의한 바와 같이 협업 필터링은 아직 해결해야 할 많은 문제점들이 존재하고

이를 극복하기 위한 노력이 지속되어야 하는 것이 사실이다. 예를 들면, 사용자들 간의 유사도를 측정하고 이용자 프로파일을 구축하고 갱신하기 위해서는 개인적 경험과 관련된 데이터를 수집하고 공유하는 것이 필수적이라 하겠다. 이와 같이, 사용자들의 개인 정보를 수집·활용하기 때문에 개인의 프라이버시 문제에 직면할 수 있다. 이러한 문제점들을 완화할 수 있는 노력을 지속함으로써 협업 필터링은 더욱 효과적으로 활용될 수 있을 것이다.

참 고 문 헌

- 김경재 · 안현철 (2009). 개인화된 추천시스템을 위한 사용자-상품 매트릭스 축약기법. <Journal of Information Technology Applications & Management>, 16(1), 97-113.
- 김지영 · 서용교 · 서길수 (2008). 상품휴기 작성자에 대해 상품휴기 독자가 느끼는 유사성이 상품휴기 독자에게 미치는 영향. <경영정보학연구>, 18(3), 67-90.
- 박태훈 · 이충재 (2014). 개인화된 영화 서비스 왓차를 개발하는 프로그램(Frograms). <정보과학회지>, 32(7), 60-62.
- 안성만 · 김인환 · 최병구 · 조운호 · 김은홍 · 김명균 (2012). 소셜네트워크 분석을 통한 협업필터링 추천 성과의 이해. <한국전자거래학회지>, 17(2), 120-147.
- 유석중 (2017). 평가 신뢰도를 고려한 협업필터링의 유사도 모델. <한국정보기술학회논문지>, 14(12), #1-146.
- 이석필 (2013). 소셜 네트워크 기반 맞춤형 콘텐츠 추천 시스템. <방송공학회논문지>, 18(1), 98-105.
- 이석환 · 박승현 (2011). 협력적 필터링 추천 시스템의 정확도 개선. <2011 대한산업공학회 춘계공동학술대회논문집>, 383-390.
- 이오준 · 유은순 (2015). 추천 시스템의 성능 안정성을 위한 예측적 군집화 기반 협업 필터링 기법. <지능정보연구>, 21(1), 119-142.
- 이종설 · 김정원 · 장달원 · 신사임 · 이명춘 · 윤경로 (2017). 소셜미디어 서비스를 위한 콘텐츠 서비스 기술. <방송과 미디어>, 22(4), 26-33.
- 정중희 · 김종우 (2013). UCC 추천을 위한 사회 네트워크 분석을 활용한 협업 필터링 기법 연구. <한국정보기술학회 논문지>, 11(1), 185-195.
- 최슬비 · 광기영 · 안현철 (2016). 사용자 간 신뢰관계 네트워크 분석을 활용한 협업필터링 알고리즘의 예측 정확도 개선. <지능정보연구>, 22(3), 113-127.
- 한소월 · 이민수 (2012). 소셜 정보 추천기법을 위한 빅데이터 모델. <정보과학회논문지: 데이터베이스>, 39(6),

- 380-386.
- Abdollahpouri, H., & Essinger, S. (2017). Multiple stakeholders in music recommender systems. *ArXiv: 1708.0012v1*.
- Bennett, J., & Lanning S. (2007). The Netflix prize. *Proceedings of KDD Cup and Workshop*.
- CACM (1997). Special issue on recommender systems. *Communications of the ACM*, 40(3), 285-295.
- Forrester Research (2008). The customer experience index.
- Haupt, J. (2010). Last.fm: People-powered online radio. *Journal of Music Reference Services Quarterly*, 12(1-2), 23-24.
- Kim, J., Oard, D. W., & Romanik, K. (2000). Using implicit feedback for user modeling in Internet and Intranet searching. *Technical Report, College of Library and Information Services, University of Maryland at College Park*.
- Konstan, J. A., & Riedl, J. (2012). Deconstructing recommender systems. *IEEE Spectrum*.
- Konstan, J. A., Miller, B. N., Maltz, D., Herlocker, J. L., Gordon, L. R., & Riedl, J. (1997). GroupLens: Applying collaborative filtering to usenet news. *Communication of the ACM*, 40(3), 77-87.
- Kumar, N., & Benbasat, I. (2006). The influence of recommendations and consumer reviews on evaluations of website. *Information Systems Research*. 17(4), 425-239.
- Miller, B. N., Albert, I., Lan, S. K., & Konstan, J. A. (2003). MovieLense unplugged: Experiences with an occasionally connected recommender system. *Proceedings of the 8th International Conference on Intelligent User Interfaces*, 263-266.
- Mudambe, S. M., & Schuff, D. (2010). What makes a helpful online review? A study of customer reviews on Amazon.com. *MIS Quarterly*. 34(1), 185-200.
- Oard, D. W. (1997). The state of the art in text filtering. *User Modeling and User-Adapted Interaction*, 7(3), 141-178.
- Sheth, B. D. (1994). *A learning approach to personalized information filtering*. Master's thesis, Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science.
- Stevens, C. (1993). *Knowledge-based assistance for accessing large, poorly structured information spaces*. Ph.D. thesis, University of Colorado, Department of Computer Science, Boulder.

Study on the Development of Collaborative Filtering Systems and Its Application

Jinmook Kim

Kangnam University

This study examined the development of collaborative filtering systems and discussed its applications for providing a customized information service. To achieve the purpose of the study, 1) the development of information filtering, including collaborative filtering, was examined; 2) various cases of collaborative filtering systems were investigated; and 3) the potential problems of a collaborative filtering system were presented. In conclusion, this paper propose several ways of applying collaborative filtering. Providing personalized information services is becoming increasingly important as artificial intelligence and IoT (Internet of Things) advance. This study will provide a theoretical basis for the design of personalized information services.

Keywords: Information Filtering, Collaborative Filtering, Personalized Information Service, Information Retrieval